Journal of
Clinical Psychology

OCTOBER 1952
Volume VIII Number 4

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Medical College, University of Vermont

JERRY W. CARTER, JR.
U. S. Public Health Service

HOWARD F. HUNT
University of Chicago

WILLIAM A. HUNT
Northwestern University

ELAINE F. KINDER
New York State Department of Mental Hygiene

C. M. LOUTTIT
University of Illinois

ROBERT I. WATSON
Washington University

ELEANOR C. THORNE, *Business Manager*

Journal of Clinical Psychology
5 Pearl Street - - Brandon, Vermont

The Brown Schools

For Exceptional Children

RESIDENT TREATMENT CENTER

Six Resident Centers in and near Austin, Texas
Administrative Offices 1110 East 32nd Street, Austin

Bert P. Brown, F.A.A.M.D., *President*

Lyndon L. Brown, *Vice-President*

Jesse J. Villarreal, Ph.D.
Speech Pathologist

Paul L. White, M.D., F.A.C.P.
Certified by American Board of Psychiatry

M. D. Heatly, M.D., F.A.C.S.
Resident Physician

H. R. Gipson, B.B.A.
Institutional Manager

L. H. Klaeveman, B.B.A., C.P.A.
Business Manager

Year-Round School

Summer Camp Program

Spring-Fed Private Pool

Address Inquiries To

Mrs. Nova Lee Dearing, *Registrar*

- **Exceptional Children Ages 4 to adulthood**
- **Comprehensive medical, psychiatric, psychologic and educational evaluation**
- **Homogeneous grouping**
- **Controlled environment**
- **Milieu approach to development of the whole child**
- **Specialized therapies**
- **Vocational training**

All the disciplines of physical and mental health in a coordinated team approach work together for the continued growth of the exceptional child

Viewbook on request

Interviews arranged with the Consulting Staff at the Schools or in your home

• An exclusive school for children with emotional or educational difficulties

JOURNAL OF CLINICAL PSYCHOLOGY

VOL. VIII OCTOBER 1952 No. 4

CONTENTS

Empirical Evaluation of Short W-B Scales. Arnold H. Hilden, Veterans Administration Regional Office, Saint Louis, and James W. Taylor and Philip H. DuBois, Washington University..... 323

The Use of the Kent EGY for the Detection of Malingering. Miriam E. Crowley, Cushing Veterans Administration Hospital..... 332

Relaxation as an Adjunct to Psychotherapy. Maja Schade, Thelma Hruza, Annette Washburne and Marie Carns, University of Wisconsin 338

On the Mathematical Representation of the Syndrome. Frank M. du Mas, Louisiana State University..... 347

The Application of an Experimental "Memory Battery" in the Psychiatric Clinic. John R. Knott, John P. Umberger and Ina McD. Bilodeau, Iowa State Psychopathic Hospital and the State University of Iowa College of Medicine..... 355

A Factor Analysis of the Bender-Gestalt Tests of Mental Patients. Wilson H. Guertin, Beatty Memorial Hospital..... 362

Relationship of Rorschach Scoring Categories to Modes of Perception Induced by Intermittent Photic Stimulation - - A Methodological Study of Perception. Ruth G. Matarazzo, Robert I. Watson and George A. Ulett, Washington University..... 368

Levels of Data in Rorschach Interpretation. Betty McD. Caldwell, George A. Ulett, Ivan N. Mensh and Samuel Granick, Washington University..... 374

Rorschach Test-Retest as a Gauge of Progress in Psychotherapy. Johanna Krout and Maurice H. Krout, Chicago Psychological Institute, and Theodore J. Dulin, Chicago Psychological Institute and Psychiatric Institute, Municipal Courts of Chicago..... 380

The Discriminability of the "Drive Factors" Represented by the Szondi Pictures. Helen M. Richardson, N. J. College for Women, Rutgers University..... 384

A Rorschach Study of the Personality Structure of Obese Women. Benjamin Kotkov, U. S. Public Health Service, and Benjamin Murawski, Boston VA Mental Hygiene Unit..... 391

The Detection of Paranoid Trends by the Eight Card Redrawing Test (8CRT). Leopold Caligor, New York University -- Bellevue Medical Center..... 397

The Effects of Two Simultaneous Cognitive and Affective Stimuli. Emanuel Starer, Veterans Administration Hospital, Coatesville, Penna..... 402

The Time-Limit Version of the Word Association Test. J. P. S. Robertson, Netherne Hospital, Coulsdon, England..... 405

(Over)

CONTENTS—Continued

Effect of Test-Retest Interval and Item Arrangement on the Shortened Forms of the MMPI. Gordon L. MacDonald, Toledo State Hospital... 408

The Consistency of Longitudinal Performance on the Graham-Kendall Memory-For-Design Test. Renate Gerborh Armstrong. Alton State Hospital... 411

Suggestions for Efficient Presentation of the Wechsler-Bellevue Object-Assembly Sub-test. Walter Shannon and Philip D. Rossi, St. Lawrence University... 413

The Ancestry of Id. Philip L. Harriman, Bucknell University... 416

Note on a Horney Hypothesis. William Seeman, Mayo Clinic... 418

Editorial Opinion... 421

Book Reviews... 424

Title Page, Volume VIII, 1952... 427

Subject Index, Volume VIII, 1952... 428

Author Index, Volume VIII, 1952... 431

GENERAL INFORMATION

The *Journal of Clinical Psychology* is an independent journal dedicated to the advancement of the clinical method in psychology. Although primarily a scientifically oriented professional journal limited to publication of original research reports and authoritative theoretical articles, it aims to foster the promotion and expansion of clinical psychology as an applied science. Special sections in each issue will be devoted to news, comments, and letters to the editors on matters pertinent to the profession. Communications and articles submitted for publication are welcome and should be addressed to the Editorial Office, 5 Pearl Street, Brandon, Vermont.

ORIGINAL ARTICLES are published only with the understanding that they are contributed exclusively to the *Journal of Clinical Psychology*. The publishers are not responsible for statements made or opinions expressed by contributors in articles printed in these columns. Copyrights cover publication in the *Journal of Clinical Psychology* and articles may not be reproduced without permission of the publishers.

MANUSCRIPTS must be typewritten, double-spaced with good margins, on one side of the paper only. The publishers will not be responsible for manuscripts lost in the mails or through unavoidable accident. Retain a carbon copy after sending the original to the editor. It is requested that all articles be as brief as consistent with clarity. All manuscripts are subject to editorial modification.

THE STYLE of all manuscripts including bibliographies must be in accordance with the instructions given in: Anderson, J. E., and Valentine, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.*, 1944, 41, 345-376. References should be arranged alphabetically according to the directions given above. Full address of the author should appear somewhere in the article, preferably at the end.

REPRINTS are furnished on order only and must be requested on the order blank provided when galley proofs are returned. Change of address notices should be sent one month before moving and should include old and new addresses. Undelivered copies will be remailed only at subscribers' expense.

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Rutland, Vermont, by Frederick C. Thorne, M.D. Subscription price $6.50; Canadian subscriptions $7.00; foreign subscriptions $7.50 including postage; single copies $2.00. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Reentered at the post office at Brandon, Vermont. Additional entry at the post office at Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont.

Copyright, 1952, Frederick C. Thorne, M. D.

EMPIRICAL EVALUATION OF SHORT W-B SCALES*

ARNOLD H. HILDEN and JAMES W. TAYLOR AND PHILIP H. DUBOIS

VA Regional Office, St. Louis *Washington University*

INTRODUCTION

Since the appearance of Rabin's article[16] on a short form of the Wechsler-Bellevue Scale, a continued interest in this subject has been shown.[1 to 15, 19]. A variety of short scales have been studied, ranging from two to five subtests, a few of them using only verbal subtests, but generally including both types. Throughout this discussion, we use the system of abbreviated notation proposed by Rabin and Guertin[17]: single letters for the verbal subtests (I, C, A, D, S, and V for subtests Information, Comprehension, Arithmetic, Digit Span, Similarities, and Vocabulary, respectively), and two-letter combinations for the performance subtests (PA, PC, BD, OA, and DS for subtests Picture Arrangement, Picture Completion, Block Designs, Object Assembly, and Digit Symbol, respectively). W-B is the abbreviation for the Wechsler-Bellevue Intelligence Scale, while VS, PS and FS stand for the Verbal Scale, the Performance Scale and the Full Scale, respectively. CS(V) refers to Hunt's short scale[8, 9, 10], in which a short form of the Stanford-Binet vocabulary test replaced the vocabulary subtest of W-B.

Short Forms Studied. The following combinations of subtests utilized as short forms have been reported: *pairs:* C and A[3, 5 8, 14]; D and PA[2, 3, 7, 8, 14]; D and DS[4]; S and BD[4]; *triads:* I, C, BD[4]; C, D, S[1]; C, A, S[2, 3, 8, 14, 16, 19]; C, S, (V) [8, 9, 10]; C, V, DS[14, 15]; *quartets:* I, S, V, BD[11]; I, PA, PC, DS[3, 14]; C, D, S, BD[2, 3, 6]; C, S, PA, BD[4]; C, V, PC, BD[14, 15]; *quintet:* C, D, A, S, PA[8], and McNemar's study[12], which reports the best pairs, triads, quartets and quintets.

Criteria Used. As a general rule, the criterion for the short forms was FS, either in terms of the weighted score or in terms of the IQ, and usually without the inclusion of subtest V[3, 4, 6, 12, 14, 16, 19]; in two exceptions[11, 15], V was included. In one study[5], the criterion used was VS IQ; another group of studies[8, 10], used two criteria, the Navy General Classification Test and a five-subtest short form which had been validated against FS; and another study[19], in addition to using FS exclusive of V, also used FS minus his short battery.

Populations. The size of the samples studied varies from 32[2] to about 500[7, 8, 10]. With one exception[12], using 355 cases from Wechsler's standardization group, the studies have been based on atypical groups of one sort or another. Among these have been court cases in a short-contact clinic[1]; institutionalized mental defectives[3, 4]; superior individuals[2]; white male recruits of questionable intellectual ability[5]; Federal prisoners, unselected, mainly psychotic, physically handicapped and otherwise maladjusted patients[6], and Federal prisoners, with some special psychometric or psychiatric reason for referral[7]; Naval recruits, randomly selected, without neuropsychiatric diagnosis[8, 10]; clinic cases referred for various emotional problems, but with no psychotics included[11]; closed ward Army hospital patients, with severe psychiatric problems, involving the major neuroses and psychoses and with several organic brain cases[14]; a mixed group, combining male veterans at a guidance center, graduate students, Army hospital cases with no psychiatric disability, and Rapaport's[18] group of state highway patrolmen[15]; female student nurses and state hospital patients selected at random from the files[16]; and Naval recruits suspected of being mentally retarded, none presenting any psychiatric problems[19].

*Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the authors are the result of their own study and do not necessarily reflect the opinion or policy of the Veterans Administration. From the Veterans Administration Regional Office, St. Louis, Missouri.

McNemar's study [12], which was not included in the above list, is the only study based on a representative group of subjects. However, it is unfortunate that the Vocabulary subtest was not included in his study, because of the absence of data on that subtest in Wechsler's manual. This is generally considered one of the most important subtests in the scale.

With regard to other aspects of the population studied, such as race and sex, the studies vary considerably, and often suffer in completeness by not reporting data on these points. As to sex, male and female are included in some [3, 4, 16]; female only in one subgroup [16]; male, or presumably male, only in the others. Wechsler does not feel that there are statistically significant differences in the total score between sexes [20, page 106]. With regard to race, of those reporting racial data, some involve white male subjects only [9, 11], and others include a mixture of whites and negroes, while one [7] included some Chinese. Where both negroes and white persons are involved, some studies report the relative proportion [5, 7], but most do not refer to it. In some instances, the author may have avoided the inclusion of negroes and made no specific mention of it, but in view of Wechsler's opinon [20, page 107] that his "norms cannot be used for the colored population of the United States," it would seem a point important enough to recognize and clarify in a report.

Age. The ages of the subjects in most reports tend to fall in the late teens or early twenties. Five studies show an average of about 23 years, with an SD of 5 to 7 years, and with a range of from late teens to close to forty years [3, 4, 5, 14, 15]. In two studies [8, 10], the average age was about 17 years, with a narrow SD of only .26 years, and with ranges of 17 to 20 and 17 to 25 years. A few studies [7, 16, 18], report only the range, and one study [17] gives no data on age, but the population consists of adults (Federal prisoners). McNemar's study was based on data for age groups 20 to 34 years.

Intelligence Level and Distribution. Most of the studies involve populations below average intelligence. Two groups [2, 4], are restricted to mental defectives and borderline cases with an average IQ of about 65. Two other studies have an average IQ in the borderline classification [11, 14], but the individuals range from mental defective to bright or superior. Another report is based on a group averaging in the dull normal classification [18], ranging from borderline to bright. In another publication [15], the mean IQ 109 is about average, with subjects ranging from borderline to very superior. The group of Naval recruits [8, 10] are close to the population at large in both mean ability and range; these studies use a short scale which is not composed entirely of Wechsler-Bellevue subtests.

The dispersion of obtained IQ's may be compared with the theoretical SD of 14.83 for the distribution of W-B IQ's. The SD's obtained on the populations with FS range from as low as 7.66, to as high as 20.66. The significance of this factor in relation to the size of the correlation coefficient does not appear to have been recognized adequately, or at least not given explicit recognition in the published reports. Only one author makes a definite comment on this point. Having observed that the size of the *r* obtained by him was lower than that obtained by another author using the same brief scale, he states that the narrow range of scores is possibly responsible, but the SD of his group is only about half a point less than that of the other group. Another author, comparing two groups of subjects, notes that the group with wider range of intelligence gave results that were "even more in agreement," i.e., produced a higher correlation between the brief scale and FS. It is difficult to know whether this indicates a recognition of restriction of range. The respective SD's of the two groups are not given.

Inspection of the results of these various studies indicates that differences in variability in the samples are not sufficient to account for the obtained differences in correlation. Other factors are also operating. There is the question of the degree of intercorrelation of the particular subtests used in any given brief scales, as well as the correlation of each subtest with the criterion, the FS. The number of subtests

used is also an important consideration, the more tests the greater the r would tend to be, other factors being equal. There is the question of whether relations differ at different intelligence and cultural levels, or of differences in functioning of particular classes of patients on different subtests and for different levels of difficulty.

Results of the Present Study

The population of our study was drawn from cases seen at a VA Mental Hygiene Clinic, most of whom are classified as psycho-neurotics. All patients who had been tested up to June, 1949, were included, with the exception of 12 cases eliminated because of race (Negro), age (above 44 or below 20 years), or intelligence level (IQ below 67). This resulted in a group that consisted of 161 white male veteran patients, none over 44 or below 20 years of age, and with a minimum IQ of 67. The restriction of age range served the purpose of limiting the subjects to World War II veterans, and also coincides with ages covered in Wechsler's standardization tables[20, Tables 39 and 40]. The cross-validation group consisted of 81 patients tested subsequent to June, 1949.

As can be seen from Table 1, both the experimental and cross-validation groups yield statistical values fairly representative of the general population. The mean Full Scale IQ and SD are close to those of Wechsler. While the mean and SD of the subtests show some deviations, they are quite close to those of Wechsler for this age range. Since Wechsler presents his subtest means and SD's in smaller age groups (20-24, 25-29, 30-34, 35-39, 40-44) than the present data, his results were combined for comparison. The following formulas were used:

$$M_x = \frac{N_a M_a + N_b M_b + N_c M_c + N_d M_d + N_e M_e}{N}$$

$$\sigma_x = \sqrt{\frac{N_a(\sigma_a^2 + M_a^2) + N_b(\sigma_b^2 + M_b^2) + N_c(\sigma_c^2 + M_c^2) + \ldots}{N} - \frac{(N_a M_a + N_b M_b + N_c M_c + \ldots)^2}{N^2}}$$

Table 1. Comparison of Experimental and Cross-Validation Groups with a Sample of Wechsler's Standardization Group (Ages 20-44)

Subtests	Experimental Group N = 161		Cross Valid. Group N = 81		Wechsler Age 20-44 N = 530	
	M	SD	M	SD	M	SD
Information	9.50	2.85	9.19	3.20	10.05	3.18
Comprehension	9.97	3.03	9.60	2.96	10.07	3.11
Digit Span	7.32	3.21	7.44	3.35	9.39	3.17
Arithmetic	8.64	3.83	8.85	4.36	9.39	3.13
Similarities	9.11	3.33	9.14	3.54	9.79	2.52
Vocabulary	8.89	2.85	8.74	2.88	....	....
Picture Arrangement	8.55	3.09	8.59	3.29	9.45	3.30
Picture Completion	9.63	3.34	9.60	3.16	9.61	3.15
Block Designs	9.11	3.17	9.79	3.69	9.77	3.24
Object Assembly	9.61	2.92	10.64	3.02	9.56	3.10
Digit Symbol	8.21	3.02	8.41	3.01	9.67	3.27
Verbal Wtd.	44.54	12.75	44.16	15.14		
Performance Wtd.	45.11	11.93	46.98	13.29		
Full Scale Wtd.	89.65	22.80	91.14	25.91		22.29
Verbal IQ	97.61	14.92	97.06	16.87		
Performance IQ	97.53	15.37	100.14	17.64	.....	.....
Full Scale IQ	97.52	15.07	98.53	17.67	100.00	14.83
Age	29.07	5.84	29.69	5.64		
Education	9.21	3.11	8.88	3.17		

In these formulas, M_x and σ_x represent the subtest mean and SD for the total age range. N_a, M_a, and σ_a represent the number, mean and SD of a subtest for a particular age range; subscripts b, c, d, and e represent the other age ranges.

The average subtest mean and SD for the experimental group, not including Vocabulary, is 8.97 and 3.28. In comparison, Wechsler's group has a mean of 9.67 and SD of 3.13. The age range used by McNemar was 20-34 and this range has an average subtest mean of 9.96, and SD of 3.01. Our means are slightly lower and our SD's slightly greater than those found with the standardization group.

Tables 2 and 3 present the intercorrelations of sub-tests and their correlations with the Full Weighted score and the Full Scale IQ. There are two additional rows

TABLE 2. TABLE OF INTERCORRELATIONS OF SUBTESTS IN EXPERIMENTAL GROUP N = 161

Subtests	I	C	D	A	S	V	PA	PC	BD	OA	DS
C	.61										
D	.43	.39									
A	.63	.57	.46								
S	.69	.66	.38	.59							
V	.81	.67	.38	.64	.71						
PA	.48	.53	.29	.40	.46	.53					
PC	.51	.48	.20	.47	.57	.60	.61				
BD	.58	.55	.34	.56	.58	.65	.52	.62			
OA	.25	.28	.03	.17	.34	.29	.35	.41	.47		
DS	.50	.50	.38	.43	.49	.56	.45	.52	.56	.31	
F Wtd.	.80	.77	.53	.74	.80	.84	.71	.76	.81	.50	.72
F IQ	.79	.77	.51	.74	.76	.84	.70	.76	.79	.49	.69
Corrected r											
F Wtd.	.79	.76	.52	.73	.80	.84	.70	.76	.81	.49	.71
F IQ	.79	.76	.51	.74	.76	.83	.69	.75	.78	.48	.68

TABLE 3. TABLE OF INTERCORRELATIONS CROSS VALIDATION GROUP N = 81

Subtests	I	C	D	A	S	V	PA	PC	BD	OA	DS
C	.61										
D	.55	.44									
A	.71	.64	.64								
S	.78	.62	.48	.64							
V	.83	.62	.60	.74	.75						
PA	.62	.48	.39	.51	.47	.58					
PC	.64	.65	.47	.59	.60	.71	.59				
BD	.61	.53	.52	.57	.59	.68	.54	.74			
OA	.41	.39	.43	.37	.42	.45	.43	.58	.64		
DS	.65	.51	.54	.58	.59	.61	.56	.58	.68	.43	
F Wtd.	.85	.74	.70	.81	.80	.87	.72	.83	.84	.65	.78
F IQ	.86	.74	.70	.83	.79	.88	.69	.83	.80	.62	.77
Corrected r.											
F Wtd.	.82	.69	.64	.76	.75	.84	.67	.79	.80	.60	.73
F IQ	.81	.67	.64	.77	.74	.84	.62	.78	.74	.55	.72

of correlations. These are the corrected correlations of the Full Wtd. and the Full IQ. As previously noted, earlier studies have utilized a wide variety of subjects. As a result, means of these populations are frequently far removed from the normal. Also the standard deviations have been much smaller in some cases and much larger in others. This makes comparison of results almost impossible. A correction can be applied to the standard deviation to estimate results which would have been ob-

tained had the sampling been more accurate. The following formula was used to correct[13, p. 126]:

$$R_{xy} = \frac{r_{xy}\frac{\Sigma_x}{\sigma_x}}{\sqrt{1-r_{xy}^2 + r_{xy}^2\frac{\Sigma_x^2}{\sigma_x^2}}}$$

R_{xy} is the corrected correlation, r_{xy} the obtained correlation, Σ_x the known standard deviation of the general population (in Wechsler's data, 14.8258 for the Full Scale IQ, and 22.2929 for the Total Weighted), and σ_x the obtained standard deviation.

This correction serves two purposes with the present data. It allows the results of the experimental and cross-validation groups to be compared more adequately, and the corrected values can be compared with correlations previously reported. While some discrepancies exist between the correlations of the subtests and the Total Weighted for the two groups, the differences do not seem to be important.

In the selection of the best combinations of subtests the Total Weighted score was used as the criterion rather than the Full Scale IQ. There are two reasons for this choice. When a combination of subtests is used, the score which will be obtained is the total weighted or an approximation of this. This approximation will then be used to enter the IQ tables for that particular age. It is logical then that the correlation with the Total Weighted is the most important, and avoids the interfering affect of age on IQ as determined by the conversion tables. These are extraneous factors to the development of a short form. Actually the correlations of the subtests with FS Weighted Score and FS IQ are in close agreement. It may be noted in addition that it would be sounder statistically to correlate the subtests with the Total Weighted minus that subtest. However, in practice we are not trying to predict what the total would be if the subtest were not included, but rather what the total is including this subtest.

In Tables 4 and 5 the ten best combinations of two and three sub-tests are listed. All combinations were calculated using the following formula, in which t stands for the Total Weighted (criterion), s for any sub-sum of subtests a+b+c+.. .. ..i.

$$r_{st} = \frac{r_{at}\sigma_a + r_{bt}\sigma_b + r_{ct}\sigma_c + r_{dt}\sigma_d}{\sqrt{\sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_d^2+2r_{ab}\sigma_a\sigma_b+2r_{ac}\sigma_a\sigma_c+2r_{ad}\sigma_a\sigma_d+2r_{bc}\sigma_b\sigma_c+2r_{bd}\sigma_b\sigma_d+2r_{cd}\sigma_c\sigma_d}}$$

TABLE 4. OBTAINED CORRELATIONS OF THE TEN BEST COMBINATIONS OF TWO TESTS WITH TOTAL WEIGHTED SCORE AND FULL IQ, TOGETHER WITH CORRELATIONS CORRECTED TO STANDARD DEVIATIONS REPORTED BY WECHSLER

Pairs of Subtests	Total Weighted				Full IQ	
	Experimental Group N = 161		Cross Valid. Group N = 81		Experimental Group N = 161	
	Obt.	Cor.	Obt.	Cor.	Obt.	Cor.
V, BD	.91	.91	.93	.91	.89	.89
S, BD	.91	.91	.92	.89	.87	.87
I, BD	.91	.90	.94	.92	.89	.89
C, BD	.90	.90	.91	.88	.88	.88
V, PC	.90	.89	.92	.90	.89	.88
I, PC	.90	.89	.93	.91	.89	.89
C, PC	.89	.89	.87	.83	.88	.88
S, V	.89	.88	.89	.86	.86	.86
S, PA	.89	.88	.89	.86	.86	.85
V, PA	.89	.88	.89	.86	.87	.87

TABLE 5. OBTAINED CORRELATIONS OF THE TEN BEST COMBINATIONS OF THREE TESTS WITH TOTAL WEIGHTED SCORE AND FULL IQ, TOGETHER WITH CORRELATIONS CORRECTED TO STANDARD DEVIATIONS REPORTED BY WECHSLER

Triads of Subtests	Total Weighted				Full IQ	
	Experimental Group N = 161		Cross Valid. Group N = 81		Experimental Group N = 161	
	Obt.	Cor.	Obt.	Cor.	Obt.	Cor.
S, PA, BD	.94	.94	.95	.93	.91	.91
S, V, BD	.94	.93	.94	.93	.91	.90
I, C, BD	.94	.93	.96	.94	.92	.92
I, C, PC	.94	.93	.93	.91	.93	.93
I, PC, BD	.94	.93	.95	.94	.92	.92
C, V, BD	.94	.93	.95	.94	.92	.92
S, BD, DS	.94	.93	.93	.91	.90	.89
V, PA, BD	.94	.93	.94	.93	.92	.91
I, PA, BD	.93	.93	.94	.93	.92	.92
C, PC, BD	.93	.93	.93	.90	.92	.92

All possible combinations of two and three subtests were calculated, and the ten best selected for presentation. Those selected were then checked by calculating the result for that combination in the cross validation group. The correlation of the experimental group with the IQ was also obtained. These three correlations for each selected combination were then corrected for standard deviation.

There are 55 possible combinations of two subtests; the range of correlations is from .911 for V and BD to .716 for D and OA. This latter combination has a lower correlation than that of most of the individual subtests. In Table 7 the present data are compared with combinations previously reported. Since McNemar used Wechsler's standardization data, a comparison with his results holds more meaning than a comparison with a number of the other studies in which the mean and SD are often very discrepant. In using all possible correlations for selecting the ten best combinations of two tests, there are three agreements with the list presented by McNemar. If V is eliminated from our combinations, there are two more agreements with McNemar.

There are 165 possible combinations of three subtests. The range of correlations is from .942 for S, PA and BD to .827 for D, OA and DS. This latter correlation is below that of V alone, .845. The overall agreement of our best ten with McNemar's is two; however, if V is eliminated there are four agreements. It is apparent that V correlates higher with Total Weighted than any of the other subtests, and also enters into the ten combinations very frequently. Since Wechsler does not present data on the V test, McNemar's selection of subtests is necessarily limited. This is not only unfortunate for comparative purposes, but also for clinical usefulness, since this subtest is one of the most valuable for obtaining qualitative material about a patient.

The correlation of S, PA, BD with T Wtd. is .942, and that of S, V, BD is .937. While the latter is slightly lower, it would be preferable for clinical usage since the substitution of V for PA reduces the correlation by a very small amount, but adds tremendously to the qualitative diagnostic value. This material may be suggestive of distorted thought processes, which are not considered in the intelligence level as such, and which could not be obtained from PA.

To obtain combinations of four tests, reported in Table 6, the four best combinations of three were used as a core, and the remaining subtests were added to this. Therefore, the results do not necessarily represent the best possible combinations of four subtests. A complete four test analysis would probably not give significantly different results. It is obvious that with more subtests, a higher correlation will be found, since the subtests are a part of the total that is being predicted. There is a

point of diminishing returns, and this appears to be with the addition of a fourth test, the increase being only .02, from .94 to .96. One might also question the purpose of a short form. If the purpose is to save time and to accomplish the best job with the smallest number of tests, three subtests give adequate results.

TABLE 6. OBTAINED CORRELATIONS OF THE SIX BEST COMBINATIONS OF FOUR TESTS WITH TOTAL WEIGHTED SCORE AND FULL IQ, TOGETHER WITH CORRELATIONS CORRECTED TO STANDARD DEVIATIONS REPORTED BY WECHSLER

Subtests	Total Weighted				Full IQ	
	Experimental Group N = 161		Cross Valid. Group N = 81		Experimental Group N = 161	
	Obt.	Cor.	Obt.	Cor.	Obt.	Cor.
I, C, PC, BD	.96	.96	.96	.95	.95	.95
S, V, PA, BD	.96	.96	.96	.95	.93	.93
S, PA, BD, DS	.96	.96	.95	.94	.93	.92
S, V, BD, DS	.95	.95	.96	.94	.92	.92
I, C, PA, BD	.95	.95	.96	.95	.94	.94
I, C, PC, DS	.95	.95	.96	.94	.94	.94

Table 7 compares previous studies with present results. Several of the reported correlations have been corrected to the SD of the standardization data. This was not possible in all cases because the obtained SD's were not reported. It was, of course, not necessary to correct any of McNemar's correlations. There is a variation in previously reported correlations from .53 to .97. When all of these combinations are calculated on the same set of data, there is very little variation. In the present data the correlations of combinations of two tests range from .911 to .716, a difference of .195; combinations of three tests have correlations ranging from .942 to .827, a difference of .115. The range decreases with the addition of more tests, and one might well conclude that it actually does not make too much difference what selection is made.

DISCUSSION

We have discussed briefly the literature on short forms of the Wechsler-Bellevue Scale. One of the difficulties encountered in attempting a satisfactory summary of these studies has been the great variation in the character of the groups studied, and particularly the lack of completeness in what is reported. Some of the obstacles found were the use of racial groups to which the norms cannot be validly applied, the use of groups whose mean level differs markedly from that of the general population, failure to take into account the variation of the distribution of ability compared with other groups and with the general population, not taking into account the fact that correlation with the full scale is a function not only of the number of tests in the brief battery but also of correlations between the individual subtests. Some of these latter relationships may be altered as the mean level of the group falls at, below, or above the ability level of the population on which the test was standardized. Some of these problems were not met in McNemar's study, since he used the data obtained on the standardization group itself. He felt that results obtained in this way would give results more valid statistically than one based on clinical populations. We feel that it is of value to carry out a similar study with some completeness on a clinical group shown to have acceptable agreement with regard to statistical values pertinent to this type of investigation. A point of further importance is that since much use is likely to be made of short forms in clinical settings, it is advisable to determine if careful study reveals that findings for a normative group also hold generally on a clinical group. In addition, McNemar's study did not include the Vocabulary subtest, one of the most important tests in the battery both from the standpoints of predictive accuracy and clinical value.

TABLE 7. OBTAINED CORRELATIONS OF COMBINATIONS PREVIOUSLY REPORTED, TOGETHER WITH THE CORRELATIONS OBTAINED FOR THESE COMBINATIONS, ALL CORRELATIONS CORRECTED, WHERE POSSIBLE, TO STANDARD DEVIATION REPORTED BY WECHSLER

Combination	Reference	Obt. r with F Wtd	Obt. r with F IQ	Cor. r with F IQ	*Present Data* Obt. r with F Wtd	*Present Data* Cor. r with F Wtd
I, S,	12. McNemar (7)	.84			.87	.86
*I, BD	12. McNemar (1)	.88			.91	.90
C, A	3. Cotzin		.53	.77	.85	.84
	14. Patterson	.85				
*C, BD	12. McNemar (2)	.88			.90	.90
C, DS	12. McNemar (5)	.85			.86	.85
D, PA	3. Cotzin		.75	.91	.77	.76
	7. Gurvitz		.90			
	14. Patterson	.82				
A, BD	12. McNemar (9)	.84			.87	.87
A, DS	12. McNemar (10)	.84			.86	.86
S, PC	12. McNemar (6)	.85			.89	.88
*S, BD	12. McNemar (3)	.88			.91	.91
S, DS	12. McNemar (4)	.86			.88	.88
BD, DS	12. McNemar (8)	.84			.87	.86
*I, C, BD	12. McNemar (4)	.91			.94	.93
I, S, BD	12. McNemar (2)	.91			.93	.93
I, BD, DS	12. McNemar (8)	.90			.93	.93
C, A, S	3. Cotzin		.59	.82	.90	.89
	14. Patterson	.89				
	16. Rabin		.93	.89		
	19. Springer		.92	.94		
C, A, BD	12. McNemar (7)	.90			.92	.91
C, S, (V)	8. Hunt	.86			.91	.90
C, S, BD	12. McNemar (9)	.90			.93	.93
C, V, DS	14. Patterson	.93			.92	.91
	15. Patterson	.90				
C, BD, DS	12. McNemar (1)	.91			.92	.92
D, S, PC	12. McNemar (10)	.90			.91	.91
D, S, BD	12. McNemar (6)	.91			.91	.90
S, PC, DS	12. McNemar (5)	.91			.92	.92
*S, BD, DS	12. McNemar (3)	.91			.94	.93
I, C, BD, DS	12. McNemar (6)	.93			.95	.95
I, S, V, BD	11. Kriegman					
	II and III		.87		.94	.94
	II plus III		.95			
I, S, BD, DS	12. McNemar (3)	.93			.95	.95
I, PA, PC, DS	3. Cotzin		.74	.91	.94	.93
	14. Patterson	.95				
C, D, S, BD	3. Cotzin		.84	.95	.93	.93
	6. Geil		.97	.94		
	14. Patterson	.94				
C, A, PC, DS	12. McNemar (4)	.93			.95	.95
C, A, BD, DS	12. McNemar (1)	.93			.94	.94
C, S, PA, BD	4. Cotzin		.86	.96	.95	.95
C, S, BD, DS	12. McNemar (2)	.93			.95	.95
C, V, PC, BD	14. Patterson	.96			.96	.95
	15. Patterson	.96				
C, PC, BD, DS	12. McNemar (9)	.93			.95	.95
D, S, PC, BD	12. McNemar (7)	.93			.95	.95
A, S, PC, DS	12. McNemar (5)	.93			.95	.95
A, S, BD, DS	12. McNemar (10)	.93			.95	.95
S, PC, BD, DS	12. McNemar (8)	.93			.95	.95

*Combinations represented in Tables 4 and 5 as being among the ten best.

( ) Numbers after references to McNemar indicate the order of his combinations.

Summary

The present study included an experimental group of 161 cases and a cross validation group of 81 cases. These groups gave results which are consistent and quite similar to the standardization group of Wechsler in level and range of ability. All possible combinations of two and three subtests were calculated and the best ten presented. Several combinations of four were also evaluated, but it seems that a point of diminishing returns is reached between a battery of three and a battery of four subtests.

The combination which seems most fruitful clinically and yields a high relationship with the Full Scale is Similarities, Vocabulary, and Block Designs, which correlated .937 with Total Weighted Score. One combination of three subtests actually correlates slightly higher, but the slight gain in prediction would be more than offset by the sacrifice in clinical value coming from loss of the Vocabulary subtest.

Bibliography

1. Bay, Margaret and Berks, Mary. Comprehension, similarities, and digit symbols of the Wechsler-Bellevue Scale used in a court clinic. *Amer. Psychologist*, 1948, 3, 365 (Abstract).
2. Cohen, B. and Steisel, I. An evaluation of the various short forms of the Wechsler-Bellevue Test. *Proc. Ia. Acad. Sci.*, 1947, 54, 221-226.
3. Cotzin, M. and Gallagher, J. J. Validity of short forms of the Wechsler-Bellevue Scale for mental defectives. *J. consult. Psychol.*, 1949, 13, 357-365.
4. Cotzin, M. and Gallagher, J. J. The Southsbury Scale: A valid abbreviated Wechsler-Bellevue for mental defectives. *J. consult. Psychol.*, 1950, 14, 358-364.
5. Cummings, S. B. A rapid method of estimating the IQ's of subnormal white adults. *J. Psychol.*, 1946, 21, 81-89.
6. Geil, G. A. A clinically useful abbreviated Wechsler-Bellevue Scale. *J. Psychol.*, 1945, 20, 101-108.
7. Gurvitz, M. S. An alternate short form of the Wechsler-Bellevue Test. *Amer. J. Orthopsychiat.*, 1945, 15, 727-732.
8. Hunt, W. A., Klebanoff, S. G., Mensh, I. N., and Williams, M. The validity of some abbreviated intelligence scales. *J. consult. Psychol.*, 1948, 12, 48-52.
9. Hunt, W. A., French, E. G., Klebanoff, S. G., Mensh, I. N., and Williams, M. The clinical possibilities of an abbreviated individual intelligence test. *J. consult. Psychol.*, 12, 1948, 171-173.
10. Hunt, W. A., French, E. G., Klebanoff, S. G., Mensh, I. N., and Williams, M. Further standardization of the CVS individual intelligence scale. *J. consult. Psychol.*, 12, 1948, 355-359.
11. Kriegman, G., and Hansen, F. W. VIBS: A short form of the Wechsler-Bellevue Intelligence Scale. *J. clin. Psychol.*, 3, 1947, 209-216.
12. McNemar, Q., On abbreviated Wechsler-Bellevue scales. *J. consult. Psychol.*, 14, 1940, 79-81.
13. McNemar, Q. *Psychological Statistics*, New York: John Wiley and Sons, Inc., 1949.
14. Patterson, C. H. A comparison of various "short forms" of the Wechsler-Bellevue Scale. *J. consult. Psychol.*, 10, 1946, 260-267.
15. Patterson, C. H. A further study of two short forms of the Wechsler-Bellevue Scale. *J. consult. Psychol.*, 12, 1948, 147-152.
16. Rabin, A. I. A short form of the Wechsler-Bellevue Test. *J. appl. Psychol.*, 27, 1943, 320-324.
17. Rabin, A. I., and Guertin, W. H. Research with the Wechsler-Bellevue Test: 1945-1950. *Psychol. Bull.*, 48, 1951, 211-248.
18. Rapaport, D. *Diagnostic Psychological Testing. Vol. I.* Chicago: The Year Book Publishers, Inc., 1945.
19. Springer, N. N. A short form of the Wechsler-Bellevue Intelligence Test as applied to naval personnel. *Amer. J. Orthopsychiat.*, 16, 1946, 341-344.
20. Wechsler, D. *The measurement of adult intelligence.* (3rd Ed.) Baltimore: Williams and Wilkins, 1944.

THE USE OF THE KENT EGY FOR THE DETECTION OF MALINGERING*

MIRIAM E. CROWLEY

Cushing Veterans Administration Hospital

Malingering, the feigning of a disability, is important to physicians and psychologists in the armed services. It is desirable to find a means of detecting feigned feeblemindedness, which we shall henceforth designate as malingering. Hunt and Older [3] describe some differences in performance on several mental tests between feebleminded men and malingering men. They offer no specific scoring stencil. Pollaczek [8] and Goldstein [2] describe keys for the detection of malingering. In Hunt and Older's study, the mean scores of the feebleminded group were below those of the malingerers. Goldstein's malingerers had mean scores significantly below the means of the control group, made up of "genuine failures." Pollaczek found no significant differences in the means of her experimental and control groups. In all these studies each group consisted of about 50 men.

The goal of the present study was to identify some specific quantitative signs, on the Kent Emergency Test, that would make possible the detection of persons who try to feign mental deficiency. The Kent was chosen because of its simplicity, and because answers on it seemed likely to yield differentiating cues.

SUBJECTS

Four groups of women were used as subjects (Ss). Letters used to identify each group appear in Table 1 together with the method of testing and the number of cases.

TABLE 1. SUBJECTS AND METHOD OF TESTING.

Key	Group	Method	N
A	College "malingerers"	Oral, individual	59
B	State School feebleminded	Oral, individual	58
C	College "malingerers"	Written, group	176
C′	College "straight"	Written, group	(176)*
D	Non-college "malingerers"	Oral, individual	11
D′	Non-college "straight"	Oral, individual	(11)

*Parentheses are used to indicate use of same Ss as in preceding group.

Group A consisted of 59 women, ranging in age from 17 to 22 years, who were selected at random from four classes in a women's college. Group B consisted of 58 feebleminded women who were residents in a state school.[1] They were matched with the college Ss for age. The mean mental age of the feebleminded Ss, based on the 1916 Stanford-Binet, was 9 years 11 months, with a range of 8-0 to 12-0. Only women who were not obviously feebleminded in appearance were used as Ss in this group.

To determine how college women normally answer the Kent questions, the test was administered as a group test to 176 students in an introductory class in psy-

*From the Clinical Psychology Division of the Neuropsychiatric Service, Cushing Veterans Administration Hospital, Framingham, Massachusetts.

[1]The writer wishes to thank Miss Mathias, psychologist at Belchertown State School, for her help in securing the Ss for this group, and the management and staff of this institution for their co-operation in making the research possible.

chology. For lack of a better word, we refer to this procedure as taking the test "straight." This group was also asked to feign feeblemindedness on the written test, so we obtained two groups from these Ss, namely C′ and C (see Table 1). These Ss were similar to group A in age and IQ. Intelligence tests were not administered to groups A and C since we could reasonably assume the IQ's of the college students to be at least at the bright normal level.

Group D was added in order to include non-college malingerers in the study. This group was made up of 6 nurses and 5 secretaries of average intelligence (M = 107; Range 97-113), who were not in college and had not attended college. Their ages ranged from 19 to 22. These 11 Ss were selected from 18 female hospital employees who were similar in age to groups A, B and C, and whose IQ's were determined by means of SRA Verbal Form (AH). Ss with IQ's beyond the middle or "normal" range of that test were not included in the test. Thus we excluded Ss who scored low enough to be considered feebleminded and those bright enough to be "college material." This group also took the test twice, as shown in Table 1, and is included as groups D and D′.

Procedure

All groups were given the original Kent Emergency Test, containing 25 questions[4]. Members of groups A, B, D and D′ were tested individually and gave answers orally. Three qualified students[2] assisted in testing group A, and the writer tested all Ss in groups B, D and D′. Groups C and C′ were tested in groups of about 30, during a psychology class. The test was mimeographed for these Ss who wrote their answers on the question sheet.

The following instructions were read when the test was given to groups A, C and D.

> "During the war, an experiment was conducted on Navy men in order to investigate the problem of malingering, that is, of pretending to have defects in order to avoid military service. The purpose was to study the different ways that intelligence test questions were answered by men who were imitating malingerers, and men who were really feebleminded. Now, we are trying to study the differences between normal women who are pretending to malinger, and really feebleminded women. I am going to give you a brief intelligence test. You are to imagine that women are being drafted and that you want to be eliminated from this draft. You have decided to try to get out of it by acting feebleminded. Try to answer the questions in the way *you* believe a mentally deficient person would. Don't forget, you are trying to appear feebleminded."

Results

Our data showed quantitative and qualitative differences between malingerers and truly feebleminded people. The differences noted here are suggested as cues for the detection of malingering.

Quantitative Differences in Mean Scores. The mean score of each of the malingering groups, (A, C and D) is lower than that of the feebleminded group (B), see Table 2. Differentiating a *group* of malingerers from a group of truly deficient Ss seems possible from the mean scores of groups. Scores of groups A and B, B and C, and B and D have a large overlap however, and identifying an individual malingerer by his total score may be possible only if it is an extremely low score.

Our data were the raw scores obtained from the total earned[3] on all 25 items of the original Kent Emergency Test. Thus we have a greater number of answers with which to make comparisons than are obtainable from Scale D, the 1941 Revision[5], which consists of only 10 questions. Age levels computable on Scale D[6] were not used because so many of the malingerers scored too low, i.e., below 10, and they could not be rated on the age scale.[4]

---

[2]The writer is indebted to Miriam Leveton, Nancy Nore and Joann Smith for their kind assistance with the testing.

[3]All scoring was done by the writer. Lewinski's[7] response list was used as an aid.

[4]Only 13 Ss in group A and 12 in C could be assigned a Kent Mental Level.

TABLE 2. MEAN RAW SCORE, STANDARD DEVIATION AND RANGE OF EACH GROUP.

Group	N	Mean	Sigma	Range
A	59	16.9	10.5	0–49
B	58	30.8	6.4	14–46
C	176	10.8	7.1	0–40
C′	(176)	54.4	4.8	36–67
D	11	17.1	11.1	5–44
D′	(11)	44.3	5.9	34–55

Differences between the mean scores of groups A and B, B and C, and B and D are statistically significant. Critical ratios are given in Table 3. The difference between the written and oral mean scores of the college Ss is significant and therefore groups A and C are not pooled and treated as a single group.

TABLE 3. CRITICAL RATIOS BETWEEN MEANS OF ALL GROUPS.

Groups	A	B	C	C′	D	D′
A	....	8.6	4.2	26.2	.05	15.7
B	....	....	20.0	25.8	4.0	6.5
C	....	....	....	6.8	1.8	17.3
C′	....	....	....	....	10.6	5.7
D	....	....	....	....	....	6.9
D′	....	....	....	....	....	....

The differences between the mean scores of the non-college Ss (D) and the college Ss (A and C) are not significant either for oral or written malingering (CRs = .05 and 1.80). It appears that intelligence, at least at the average and above-average levels, has no effect on the feigning of feeblemindedness.

The means obtained when the test was taken "straight" are higher than all other means, and significantly so.

Levels of Answering Questions. One might ask whether the low scores obtained by our malingering groups were not the result of frequent refusals to answer, or whether the two groups actually attempted the same number of answers and simply answered differently. We therefore counted, in groups A and B, the number of questions to which an answer was attempted, and the number to which no answer or "I don't know" was given. We also noted the number of answers which received full credit and the number receiving zero credit.

The number closest to the median of the combined distribution of groups A and B was used as a "cutting point" on each count. Frequencies below and above these points were used in a two-fold table and the chi square test, suggested by Cronbach [1] for non-normally distributed data, was used to test the null hypothesis. The median values of each group are shown in Table 4, together with the probabilities that the X^2 values obtained in the comparisons are significant. The differences between groups A and B are not significant for the number of items attempted, nor for the number not attempted. Therefore, the two groups gave answers with equal fre-

TABLE 4. SIGNIFICANCE OF DIFFERENCE BETWEEN GROUPS A AND B ON FOUR DIFFERENT COUNTS OF 25 KENT ITEMS. BASED ON X^2, DF = 1

COUNT	GROUPS		P
	A Median	B Median	
(1) Items attempted	24	24	> .95
(2) Items answered, "I don't know"	1	1	> .95
(3) Items receiving zero credit	15	9	< .001
(4) Items receiving full credit	4	9	< .001

quency, and the difference in their mean scores must be attributed to a difference in the number of correct answers, or to the difference in the value of the answers given, or both. We can reject the null hypothesis on the basis of the P values found for counts 3 and 4 (see Table 4). The two groups differ significantly in the number of items on which they obtained zero credit and full credit. Group A had more items given zero, while group B had more given full credit.

Item Analysis. Different amounts of credit ranging from zero to four are given on the Kent test. A check was made to discover on which items the feebleminded and malingerers earned different amounts. The chi square technique was applied to determine the significance of the difference in the distribution of credits earned on each of the 25 items. On all but five questions (11, 12, 16, 17 and 21), the differences in the distribution of credits were considered to be significant.[5] Except on Q. 23 and Q. 25, which are relatively difficult, the feebleminded Ss usually earned the high credits, and the malingerers received the low.

P values between 2% and 5% are taken to indicate only a tendency on the part of the groups to differ in the distribution of credits on each item. Q. 12 and Q. 17 had P values at this level.

A more critical discrimination between groups seems possible from this item analysis than from mean total scores. One way to recognize the individual malingerer is to note the questions on which he has earned very low credits. The more often he scores low on those 18 items on which most of our feebleminded Ss scored high, the more he may be suspected of malingering. Also if he does well on Q. 23 and Q. 25, but has a very low total score, he can be suspected of malingering.

Examination of credit distribution by group D malingerers revealed a scoring pattern very similar to that of group A, but group D is so small that the distribution of credits earned by it was not put to a statistical test. On six of the 20 differentiating items not a single feebleminded person failed, i.e., received zero credit.[6] Ss in groups A and B were sorted into two categories, those having zero and those having more than zero on each of these items. The differences were all significant beyond the .001 level[7]. Group B had only two failures on Q. 7 and three on Q. 9, so on these eight questions there were only five failures out of the 464 possible answers. On these eight questions, there were 188 failures among group A's answers. These eight items may be most useful when the Kent is employed to identify groups or individuals suspected of malingering, because these are items which feebleminded people seem unlikely to fail. Other studies [2, 3, 8] also report this tendency for malingerers to fail easy items and to pass hard ones, while the control Ss do the opposite. Also, the items are useful because their simplicity makes failures easily recognizable as "forced" errors.

[5] P values of 2% or less. The .001 level of confidence was found for 14 of the 20 items.
[6] Questions 1, 6, 8, 10, 14, 15.
[7] Based on X^2, d.f. = 1.

Qualitative Differences. Perhaps the most striking difference between the malingerer groups and the truly feebleminded is seen in the *kinds* of answers given. Distinct qualitative differences were apparent in both correct and incorrect answers. Thus we have an additional cue for the detection of malingering. We have tried to summarize certain response characteristics of the malingerers and we give here a few examples:

1. The malingerers' answers were often foolish to the point of being nonsensical. The question, "If your shadow points to the northeast, where is the sun?" was answered with such replies as, "In the sky" and "Under the oak tree." The feebleminded Ss all named a direction although none got it right. To the item: "Tell me the names of some birds," group A Ss often named things which are not birds at all, e.g., "mosquitoes," "airplanes," "rabbits," and "cocker spaniel." The Ss in group B named only commonly known birds.

2. The malingerers were inclined to give "wild" or exaggerated answers. On the question, "How many days in a week?" 26 failed, and gave wrong answers ranging from 2 to 82. Only three in group B failed, and they answered with "Five" or "Ten."

3. The malingerers showed a general knowledge, sophistication and humor which the feebleminded person lacks.

Q. 4 "What is sand used for?" Typical of group A answers were "hourglasses" "spinach." Feebleminded Ss rarely failed this, and those who did gave answers like "For the soil," and "To fill holes."

Q. 10 "What does ice become when it melts?" Typical malingering responses included such implausible answers as "Rocks" and the somewhat more understandable "Ginger ale." None of the feebleminded failed this question.

Q. 17 "Give me the names of some large cities." Malingerers named foreign cities and countries, Cape Cod, the United States and the Bronx. The feebleminded erred by naming states and small towns.

4. The malingerers displayed a vocabulary superior to that of feebleminded persons. On the question, "What metal is attracted by a magnet? practically every feebleminded S asked the meaning of the word "magnet" while only a few of the malingerers did so. Throughout the test, group A members used such words as "metropolis," "magnetic," "electrocute," "artificial," "asbestos," "porous," and "contradictory."

5. The malingerers answered more quickly than the feebleminded on the hard questions and more slowly on the easy ones. While we have no reaction time figures to prove this statement, the subjective experience while testing was clear enough to make this a distinct possibility.

Discussion

In view of our results, one may ask whether our malingerers knew the difference between feeblemindedness and "insanity." It would appear that they were trying to be "crazy" rather than "dumb." Some explanation at the start might have helped their understanding of the difference but that would have been tantamount to telling them how to act. Our results indicate that not only does the average person not know how to feign mental deficiency, but neither does a college student enrolled in psychology.

Hunt and Older[3] reported that the Navy men did not play "dumb" enough, but our women Ss played too "dumb." Perhaps women performing in an experiment for a teacher are not comparable with Navy men participating in an experiment under the command of an officer.

People who are merely feigning malingering may be more self assured and more easy going than those for whom the test results are crucial. Any experiment puts the

malingerer in an artificial situation, but the main task in this situation is to feign mental deficiency and that is what the true malingerer wants to do. Pollaczek[8] discusses this point and concludes that "there is little reason to suspect that any other personality or intellectual factors are operating in any important fashion to distinguish the behavior of the two types of malingerers." It seems to us that personality factors provide the best explanation for the test behavior of our malingerers. Possible explanations include: (1) the women had the chance to act stupid without fear of criticism or punishment. (2) They enjoyed playing the role of the "happy little moron," who is childlike, amusing and lovable. (3) They were able to pull jokes on the examiner even though they were following instructions. (4) They could give wrong answers and still show how clever they actually were. A deeper analysis of each of these "explanations" could of course be made to examine the real personality factors involved.

Probably it is easier to think up funny answers when writing than when talking, and distance from the examiner may provide the courage needed to give such answers. The fact that group C obtained significantly lower scores than group A supports this idea. Further support is found in the fact that the no-response total was smaller for group C than for group A.

Summary

This study located quantitative and qualitative differences between the test performance of truly feebleminded women and women who were asked to feign mental deficiency. Differences between groups were noted for total raw scores and for individual items. The full Kent EGY was used, but the comparisons made here should be applicable to many other tests, and are suggested as clues for the detection of malingering.

The results of this experiment include:

1. The mean scores of the malingering groups were significantly lower than that of the truly feebleminded group.
2. It is impossible to identify malingerers from individual total scores unless they are extremely low.
3. Malingerers were inclined to fail the easy questions and pass some of the hard ones.
4. Malingering and feebleminded groups were found to have significant differences in the distribution of scores on 20 out of 25 questions.
5. Qualitative differences in answers were apparent throughout the test.

The quantitative and qualitative differences observed here may be useful in the detection of malingering by use of the Kent or some other available tests. These differences, and possible further exploration, seem to provide a foundation for the construction of a test specifically designed for the detection of malingering.

Bibliography

1. Cronbach, L. J. Statistical methods applied to Rorschach scores. *Psychol. Bull.*, 1949, 46, 393-429.
2. Goldstein, H. A malingering key for mental tests. *Psychol. Bull.*, 1945, 42, 104-118.
3. Hunt, W. A. & Older, H. J. Detection of malingering through psychometric tests. *Nav. Med. Bull.*, Wash., 1943, 41, 1318-1323.
4. Kent, Grace H. Oral test for emergency use in clinics. *Ment. Meas. Monogr.*, 1932, No. 9.
5. Kent, Grace H. Emergency battery of one-minute tests. *J. Psychol.*, 1942, 13, 141-164.
6. Kent, Grace H. *Series of Emergency Scales. Manual.* Psychological Corporation, New York: 1946.
7. Lewinski, R. J. A qualitative analysis of the Kent Oral Emergency Test as a clinical instrument in the examination of recruits. *J. appl. Psychol.*, 1942, 26, 316-331.
8. Pollaczek, Penelope P. A study of malingering on the CVS abbreviated individual intelligence scale. *J. clin. Psychol.*, 1952, 8, 75-81.

RELAXATION AS AN ADJUNCT TO PSYCHOTHERAPY

MAJA SCHADE, THELMA HRUZA, ANNETTE WASHBURNE, AND MARIE CARNS

University of Wisconsin

Introduction

This study was undertaken in order to discover (1) if the use of a specific technique of relaxation as an adjunct to psychotherapy is effective in the relief of certain somatic and psychological manifestations and (2) if so, which types of psychiatric disorders respond best.

There have been a number of reports on this subject. Hadley,[1] Pascal,[2] Schultz,[3] and Yates[4] reported relaxation successful in alleviating signs and symptoms of tension; in promoting insight in psychosomatic relationships, and thus assisting in facing and meeting problems. Though their methods varied, these investigators observed that the reduction of neuromuscular hypertension supported the psychotherapeutic approach to the patient as well as giving him relief. In these instances the method was taught by the therapist himself and directly applied in the therapeutic session. Bagchi,[5] Behrend and Weiss,[6] Bouman *et al*,[7] Jacobson,[8] and Hosskisson,[9] have reported the alleviation of somatic and psychological symptoms through relaxation alone.

When psychosomatic disturbances are caused by deep seated personality disorders, however, it seems doubtful that lasting effects may be achieved by relaxation alone. On the other hand, the authors felt that a specific training in a type of relaxation which gave the individual a kinesthetic awareness and an intellectual means of controlling tension would be a desirable adjunct to the more traditional forms of psychotherapy. The difficulties of the patient would thus be approached through both somatic and psychological channels. The technique[10] used in this study appeared to have the rationale desired.

Method

This investigation was carried out as a joint project by members of two departments of a University. The instruction in relaxation was given in the department of physical education for women. The subjects were referred by members of the staff of the neuropsychiatric division of the student health department. No direct use of relaxation was made during the psychiatric conferences.

Subjects. The subjects of this study were eighty-two university students chosen from approximately three thousand students who sought aid in the neuropsychiatric service of the student health department during a three year period. Forty-five were females and thirty-seven were males. The age range was eighteen to forty years. Each student was given a detailed psychiatric study by a psychiatrist or a clinical psychologist and was considered a suitable case for this investigation because of the degree of neuromuscular hypertension present. Every reasonable effort was made either to establish or to rule out any organic pathology which might directly account for the signs and symptoms. In two cases, one of active duodenal ulcer and one of spastic cerebral diplegia, the organic pathology undoubtedly contributed directly to the somatic manifestations. These two cases are presented in some detail in the section entitled "Results".

Table 1 presents a summary of data considered significant from the medical point of view and for the most part is self-explanatory. It should be pointed out, however, that the somatic and psychological symptoms listed are those volunteered by the students during medical and psychiatric conferences. Symptoms so listed are often more significant, at least in the mind of the patient, than answers to check lists or responses to direct questions which might be suggestive.

Procedure. The method used in this study was designed to provide the subject with a technique for inducing within himself a condition of psychosomatic quiescence.

Through concentration on the release of muscular tension voluntary control over tension may be achieved, not only for general reduction of tension during rest, but for application in situations where excess tensions interfere with adequate adjustment.

TABLE 1. SUMMARY OF MEDICAL DATA

	Females	Males	Total	%
SOMATIC MANIFESTATIONS				
Excessive Fatigue	35	29	64	78.0
Insomnia	32	28	60	73.2
Headache	22	21	43	52.4
Aches or Pains (Vague)	20	20	40	48.8
Tremor	20	26	46	56.1
Dizziness	14	14	28	34.1
Urinary Frequency	8	10	18	22.0
Nausea	8	12	20	24.4
Constipation	11	7	18	22.0
Recurrent Diarrhea	14	6	10	12.2
Precordial Pain	2	5	7	8.5
Emesis	3	4	7	8.5
Sweating	3	11	14	17.1
PSYCHOLOGIC MANIFESTATIONS				
Confusional Thinking	26	27	53	64.6
Worry	37	28	65	79.3
Retardation	30	17	47	57.3
Weeping	25	6	31	37.8
Depression	33	20	53	64.6
Blocking	25	20	45	54.9
Intellectual Inadequacy	13	8	21	25.6
Social Inadequacy	24	25	49	59.8
Irritability	18	17	35	42.7
Dissociation	14	13	27	32.9
Phobias	7	12	19	23.2
DIAGNOSIS				
Anxiety Tension State	26	14	40	48.8
Anxiety Neurosis	7	4	11	13.4
Psychopathic Personality	3	4	7	8.3
Benign Depression	2	1	3	3.7
Cyclothymia	1	0	1	1.2
Conversion Hysteria	0	1	1	1.2
Hysteria	3	2	5	6.1
Obsessive Compulsive Neurosis	3	9	12	14.6
Paranoid Personality (Post-Psychotic)	0	1	1	1.2
Spastic Cerebral Diplegia	0	1	1	1.2

Since concentration implies effort, which is contrary to relaxation, the actual muscular release is preceded by the establishment of an attitude of quiet and passivity. This attitude is obtained by breathing exercises. The subject is told to observe his respiration without interfering volitionally with its rhythm, but rather to "let it happen" naturally. Next, the physiologically relatively passive expiratory phase is extended by "giving in" to it intentionally, thus directing it without effort. The result of these practices is a quiet relaxed breathing rhythm and a mental attitude of passivity.

The release of muscular contraction is achieved by thinking of or imagining the body part concerned in its relaxed state, i.e., an arm feels heavy, the shoulder girdle wide, etc. Since exhalation appears to facilitate relaxation,[3, p. 75; 4, p. 226] it is associated with the release of tension by the instruction to "let the arm (for example) become heavy" with the exhalation. This strong association of the expiration with the muscular release enables the individual to use the relaxed exhalation to initiate a reduction of tension on those occasions where excess tensions interfere with adequate functioning.

The training proceeds from the arms and shoulder girdle to the legs and hips and finally to the throat and face. Each step is taken when the preceding one has been accomplished and when some sensation of release has been noticed in the succeeding one. The spreading of relaxation from one part of the body to another has been observed by Schultz [3, p. 30] and it seems to happen in a fairly regular way. Waiting for the spontaneous spread seems to reduce concentrative effort which may interfere with progress. Moreover, progressively larger portions of the body are relaxed simultaneously, which further reduces concentration on details and over-all general relaxation is reached more quickly. Each practice trial is terminated by a deep inhalation and a stretching of the entire body. This aids in the re-establishment of the alert state and prevents the persistence of drowsiness.

A minimum of two, preferably three, daily practice periods is suggested, which should at first not exceed ten minutes. After two to three weeks of instruction the time limit may be extended. It appears important to hold to short practice periods, because tension results from prolonged effort, or concentration fails, and either interferes with progress.

The training begins in the supine position and is continued in the sitting position. At this stage the practice of partial relaxation outside of the regular home practice is urged. The instructions are to exhale in the relaxed manner and relax the shoulder girdle. Sufficient practice enables the individual to utilize relaxation during tension producing situations and thus assume a relaxed attitude during activity.

In this study the instruction periods of approximately 40 minutes were held twice weekly during the first week. In a few cases of very slow progress these were continued for one or two weeks longer. After that the instruction proceeded with one session a week. The average time found necessary to acquire the technique was eight to ten weeks. However, with subsequent practice, the subject may be expected to gain in depth of relaxation as well as in speed with which the relaxed state is established. The total time under instruction of the subjects of this study varied considerably, because of such factors as interference from the University schedule, the large range of disorders and conditions studied, and the individual's dependence on the instructor's help.

Results

The data on the results of this study consisted (1) of the critical evaluation of specialists, based on detailed case histories and examinations, (2) of records on the course and progress of training in relaxation, and (3) of reports of the subjects on benefits obtained from relaxation. No controls were attempted because it was felt that individuals with such diverse personalities and psychiatric disorders could not be accurately cross matched.

The general status of the subject was evaluated by the therapist at the last psychiatric conference. Sixty-eight (82.9%) of the group were judged improved; ten (12.2%) were unimproved. In four subjects (4.9%) insufficient data prevented evaluation. What portion of these improvements is the result of the relaxation training and what portion the result of the concomitant psychotherapy cannot be objectively determined. In a few cases, where psychotherapy was discontinued during the training period, relaxation has clearly played a considerable role. In other instances improvement of the general status of the subject took place while relaxation proficiency remained poor. These cases are reviewed in some detail in the reports on diagnostic groups.

Judgment of relaxation proficiency was passed on a purely subjective basis from such physical signs as general flaccidity of the skeletal musculature, quiet relaxed breathing and absence of movement of the eyelids and eyes; also on the students' report on their ability to use the technique in the control of tension. Those who were rated "excellent" (15.9%) were able to relax quickly and deeply. They were consistent in obtaining good results from practice as well as from use of the

TABLE 2. SUMMARY OF RESULTS

Diagnosis	Number	Therapist's Estimate			Student's Estimate			Relaxation Proficiency				Improvement Reported By Students				
		Improved	Unimproved	Insufficient Data	Improved	Unimproved	Insufficient Data	Excellent	Good	Fair	Poor	Somatic Manifestations	Recuperation	Generally More Relaxed	Mental Efficiency	Control of Situations
Anxiety-Tension State	40	38	1	1	37	1	2	8	21	10	1	23	18	13	6	24
Anxiety Neurosis	11	6	4	1	5	3	3	1	3	3	4	5	6	3		3
Psychopathic Personality	7	4	2	1	3	1	3	1	1	4	1	2	5	2	5	2
Benign Depression	3	2	1		2	1			2		1	2	1		1	1
Cyclothymia	1	1			1				1			1	1			
Conversion Hysteria	1	1			1				1			1			1	
Hysteria	5	3	2		4		1	1	1	2	1	4	2	3	1	
Obsessive-Compulsive Neurosis	12	11		1	11		1	1	3	4	4	4	5	3	4	7
Paranoid Personality	1	1			1				1				1	1		
Spastic Cerebral Diplegia	1	1			1			1				1				
Total Number	82	68	10	4	66	6	10	13	34	23	12					
Percentage		82.9	12.2	4.9	80.5	7.3	12.2	15.9	41.5	28.0	14.6					

technique in tension producing situations. Subjects who were rated "good" (41.5%) did not relax as rapidly, or were less consistent in reaching their maximum depth. They were successful in the application of the technique outside of practice periods. Subjects were rated "fair" (28.0%) when relaxation seemed to remain "shallow", and when success in the use of the technique was only partial or inconsistent. The rating "poor" (14.6%) was given when practice resulted in partial or occasional success only.

Table 2 presents a summary of results with the eighty-two subjects divided into diagnostic groups. The last columns give an account of alleviations of tension manifestations considered by the students as contributing significantly to the improvement of their general status and to their adjustment to University life. The figures are conservative in that they were volunteered to the instructor of relaxation.

Somatic Signs and Symptoms. Among the somatic symptoms and signs, insomnia was the one showing the most rapid improvement. Not infrequently, improvement was reported in the first two weeks of training. Only one out of sixty subjects with this complaint did not experience improvement. Headaches and abdominal pain and discomfort responded well. In some cases they gradually disappeared as the subjects became more relaxed generally. In other instances, relaxation was deliberately used to combat these symptoms.

One of the most common complaints was painful tension in the muscles of the neck and shoulders. When these were severe and of long standing it proved helpful to loosen the muscles mechanically by passive movement and stretching of the neck by the instructor. They disappeared or improved greatly during the training.

The amount of recuperation from short practice periods of relaxation is surprising to most subjects. They reported they felt more refreshed than after sleeping, particularly sleeping in the afternoon.

Psychological Signs and Symptoms. Students' reports on improved mental efficiency refer to the aid relaxation has given to study. Concentration improved and the span of attention was considerably extended by short practices of the technique while studying. In some instances the effect of the practices was, of course, a reduction of fatigue. But more often the subjects felt that inability to concentrate and a general restlessness prevented successful continuation of work, and these were overcome by relaxation.

The subjects who reported they felt generally more relaxed were those who had been conscious of and disturbed by a sense of hurry, restlessness and irritability. In some the improvement came quite automatically as their proficiency in relaxation grew. Others deliberately slowed down their pace during activity. They also became sensitive to tension and effort exerted in all kinds of activity and deliberately attempted to preserve their energy. In that way and with the help of relaxed breathing, those who had to work with great speed, as when waiting on tables or working in kitchens of large restaurants, learned to remain relatively relaxed in spite of the pressure under which they worked.

Perhaps the most significant achievement for a great number of subjects was the control of emotionally produced tensions in situations in which adjustment and adequate functioning were extremely difficult or impossible. Some practiced relaxation before the anticipated critical occasion and the after effect carried them through without further difficulties. Others used the partial relaxation shortly before or during the time of stress. Still others observed in themselves at such times localized muscular tensions, as in the shoulders, abdominal muscles or the hands and feet and found that the ability to release them combined with the relaxed exhalation, brought about a general lessening of tension. After-reactions, either somatic or psychological, ceased to occur or their duration was greatly reduced. Many of the reports of improvement in insomnia, tremor and abortion of headaches and abdominal pain belong in this category.

Diagnostic Groups

The discrepancy in numbers of the different diagnostic groups makes statistical evaluation impossible. It appears of interest then to determine the possible significance of such training in relation to the individual's difficulties and also if any particular type of case responded more or less favorably.

Anxiety-Tension State. As a group the anxiety-tension states responded well to the training in relaxation. Thirty-eight of them were estimated as improved, one as unimproved. In one case, insufficient data made estimation of the status impossible. The large majority of these subjects possessed a drive for self-efficiency. Willing to work for their own progress, they recognized and appreciated relaxation as an active control method with which they could rationally and concretely combat their difficulties. Good results seem predominantly dependent upon the individual's ability and willingness to persevere in practicing. Twenty nine of the forty subjects developed good or excellent proficiency in relaxation. Ten remained fair. Of these, four did not finish the training. Four were dependent individuals unwilling to take responsibility for themselves. They relaxed quite well under supervision and when circumstances were favorable, but never really tried to develop skill. In these the results of practice were mainly recuperation and improvement of insomnia.

Only fair relaxation was achieved by the two subjects who progressed to more serious disorders. One showed temporary improvement. He practiced relaxation regularly, reducing fatigue and feeling generally more relaxed. But under great strain, which developed from certain situational stimuli, his attempts to use the technique were only partially successful; though on less critical occasions, as examinations, it helped him considerably. A few months later he developed a schizophrenic episode and left the University. In this case relaxation could be considered a supportive measure. The second developed a schizophrenic episode seven months after termination of the relaxation training. Upon reentering the University she returned voluntarily for further instruction. After two months her relaxation proficiency became good. An abdominal discomfort could be controlled by relaxation practice and she used the technique successfully for the reduction of fatigue and general tension. This subject, though unable to relax during the acute phase of her illness, received valuable help through it during a remission.

One case of anxiety tension, judged improved by the therapist, remained definitely poor in her ability to relax. She did not practice regularly and made no attempt to do so in the face of difficulty. Considering her poor achievement, relaxation can hardly have contributed to her general improvement.

The subject with active duodenal ulcer, referred to previously, was classified in this group. He expressed great apprehension over his condition and showed signs of marked anxiety. Preparing for his Ph.D. oral examination, he insisted he could not follow the recommendation to enter the student infirmary for bed rest. After four weeks of intensive psychotherapy he was referred for training in relaxation. He grasped the essentials of the technique quickly and made rapid progress. After the second month he entered the infirmary for bed rest, relaxation instruction being continued there. He improved rapidly and was able to complete his academic work on schedule. In his last conference at the end of the academic year he stated that he continued to use the relaxation technique and felt that he had been able to accomplish more work with less expenditure of energy and much less tension. Four months later, after holding a responsible research position for three months he reported no return of symptoms.

In an attempt to evaluate relaxation as an adjunct to psychotherapy four subjects should be considered separately. These students did not receive psychotherapy. In three, initial psychotherapy had failed to win a basic rapport. They remained negativistic and unwilling to face their problems. Through relaxation they obtained relief from their tension manifestations to the extent that they were considered improved. The fourth case receiving no psychotherapy cancelled her psychiatric conferences shortly after she began the relaxation training. She developed

considerable control over her symptoms and as a result gained in self-confidence and felt much improved. She was not judged improved, however, by the therapist, since her basic anxieties were not considered relieved.

Anxiety Neurosis. In the group of eleven cases of anxiety neurosis only four followed through with both psychotherapy and relaxation. These four possessed a strong drive for self-efficiency. They developed good skill in relaxation and made intelligent use of it. All improved greatly. Two years after the training they remained improved and continued to use the relaxation technique.

One subject did not continue psychotherapy after alleviation of symptoms through relaxation and therefore no data could be obtained by the therapist. She learned to relax quickly and deeply, practicing regularly and frequently. After two and a half months she believed herself on a much better functioning level. Two other subjects were estimated by the therapist as improved. Neither completed the training. Four subjects failed to improve. They did not follow through with psychotherapy or relaxation.

Psychopathic Personality. With one exception the work with relaxation in this group proved either unsatisfactory or very tedious and prolonged. The subject with excellent relaxation proficiency had instruction for three and a half months and continued with occasional follow-up for four months more. A later written communication from her told of success and progress in work and personal adjustment. Nothing is known of the others. For these relaxation proved to be a supportive measure during the time they were under instruction. They used it for the reduction of fatigue and to increase their ability to study longer and more effectively, thus keeping them on a better functioning level.

Benign Depression. This diagnosis was applied to those patients who displayed evidence of a depressive reaction of a non-cyclic type. Etiologic factors were found in situational stimuli.

Of the three students with benign depression two improved. In one of these relaxation could not have contributed, since he remained poor in his ability to relax due to lack of practice and interest. In the other very good results were obtained. She was considered an emergency case because of retardation, dissociation, feelings of futility and suicidal ideas. After two months she relaxed well and stated she had "built the practice of relaxation into her life". A follow-up a month later showed that she had continued to make use of the technique and stated, "I simply have no reason any more to go into a panic". Written communications during a three year period state that she has had no recurrence of symptoms. The one whom the therapist did not consider to be improved was a Siamese woman whose main problem seemed to be adjustment to life in this country. She learned to relax well and reported good results from relaxation in recuperation and greater ability to concentrate. She did not return to the Clinic during the remainder of her attendance at the University.

Cyclothymia. This diagnosis was applied to one individual whose mood shifts were cyclic, often seasonal and reactive in nature. In no instance did the degree of "ups" and "downs" exceed a pattern of mild hypomania or retardation. This young woman was a rigidly tense individual. She began the training while a depression was subsiding and though progress was initially slow, she did well ultimately. She had been keenly aware of her tenseness and her difficulty in controlling it. The realization that she could learn to relax greatly increased her self-confidence and gave her considerable encouragement.

Conversion Hysteria. This diagnosis was made in one student who demonstrated transitory episodes of anxiety or panic reactions. After eight months of psychotherapy this subject had gained good insight into the basic etiology of his difficulties. He was free from all symptoms except for occasional periods of cramping and weakness of the legs. At this point he was referred for training in relaxation. After two months of instruction he relaxed deeply and quickly. When last seen he said "The leg symptoms are essentially gone; when any feeling of weakness or spasm occurs,

it can be relieved by relaxation". He has had no recurrence of his original difficulties over a three year period.

Hysteria. Only two of these subjects really applied themselves to the learning of the technique. Both improved and made good use of it. The three others never seemed to understand the procedure or how to practice it. Of these, two remained unimproved and the third, though considered improved, relaxed poorly.

Obsessive-Compulsive Neurosis. The improvement for the group was considerable. However, in four students the relaxation training could not have contributed to it, since two discontinued the instruction and all four relaxed poorly and failed to practice regularly.

Two subjects diagnosed as obsessive-compulsive neurosis with schizoid personality had considerable difficulty relaxing during periods of anxiety. They always practiced during remissions. As a result they felt more relaxed throughout the day and less fatigued. In one subject the periods of remission lengthened considerably during the two years following instruction in relaxation. Whether or not this is a result of his ability to relax must remain a matter of speculation.

Of special interest is a case of agorophobia in a female student who was unable to leave her home unless accompanied by her husband. She had marked suicidal tendencies which reached panic levels. She was also disturbed by daily episodes of feelings of unreality. She learned to relax excellently and worked with considerable tenacity on the control of her difficulties. She claimed she could control the panic when, at the first sign of it she could lie down and relax. By controlling the panic she felt the suicidal fears were eliminated. In the same way she has eliminated the feelings of unreality and has had no recurrence for the last two years. In general she felt much better, although her phobia was essentially unchanged, except for a slight extension of her radius of movement to a few blocks around her home.

Paranoid Personality (Post Psychotic). One patient could not be placed under any established diagnosis. When first seen by us he was still convalescing from a severe schizophrenic psychosis. He had had intensive treatment elsewhere with insulin and electric shock therapy. With this student regular relaxation practice under supervision was continued for a considerably longer period than with any other subject. Ultimately he relaxed well. For him the main benefit lay in the reduction of fatigue and the elimination of insomnia. The practice also reduced his tendency toward agitation and irritability.

Spastic Cerebral Diplegia. This case, referred to previously, had been observed for approximately two years in the Neuropsychiatric Clinic at infrequent intervals (nine conferences). Neurological examination revealed gross horizontal nystagmus with right homonymus hemianopsia, a right peripheral facial palsy, constant gross tremor of both hands with choreiform movements, a moderately ataxic gait and a speech defect. He was granted a graduate assistantship, which position entailed teaching with the necessity for blackboard writing. Under this stimulus the tremor and choreiform movements in his arms increased. It was felt that instruction in relaxation might help him. During the first two weeks involuntary movements in the left arm and foot were pronounced and quite distracting. With an increase in ability to relax involuntary movements disappeared. At the conclusion of the training, which lasted two and a half months, he gained considerable skill which he not only practiced regularly but which he could also apply at times of strain. The results were expressed by the student as follows: "I feel my arms are under better control; in social situations I believe the tremor is hardly noticeable; with conscious effort at relaxation I can always control my arms". As a consequence of the improvement he had gained greatly in self-confidence and poise.

Summary and Conclusions

In this study an attempt was made: (1) to evaluate a specific technique of relaxation as an adjunct to psychotherapy, and (2) to discover whether any particular types of cases responded more or less favorably to a training in relaxation.

The technique under investigation demonstrated its value to psychotherapy by the following contributions:

1. Relaxation breaks into signs and symptoms of tension, giving relief and raising self-confidence.
2. It demonstrates psychosomatic relationships to the subject. A recognition of these relationships facilitates intellectual insight and provides the possibility for rational management of tension manifestations. When medication such as sedatives is given, the individual is merely the recipient, while in relaxation he is provided with an active control method.
3. Relaxation may break up habit patterns of hurry and excessive activity thereby improving functional balance.
4. With the necessity for perseverance and self-discipline, the training in relaxation becomes a process of self-education.

As to the question which type of cases responded best to the relaxation training, certain general conclusions can be drawn. It was observed that some individuals with more serious disorders responded as well, if not better, than some with benign disorders. Hence, it is believed that response is dependent primarily upon personal qualities. Our observations lead to the following conclusions:

1. When personal qualities, such as perseverance, aggressive drives, and the will for self-efficiency are present, good and excellent results from the relaxation training can be expected even under difficult conditions.
2. The anxiety-tension cases, i.e., disorders of relatively recent occurrence, or of reactive rather than affective type appeared to respond most readily. In this relatively benign disorder the large majority had a strong desire for self-efficiency, and they were willing and able to work for their own progress.
3. Individuals who lack self-direction, either by reason of their illness or who are overly dependent do not respond favorably. For these, however, the training may constitute a supportive measure of considerable help.
4. The reduction of tension manifestations and the ability to maintain an adequate functioning level through relaxation may occasionally interfere with the individual's continuing psychotherapy. In reactive disorders the training alone may conceivably suffice because of its self-education, its breaking of habit patterns and the fact that it is an active control method. This cannot be expected for the more severe and deep seated personality disorders.

Psychotherapy depends to a large extent upon self-direction and the will for self-efficiency. These are stressed and encouraged through the training in relaxation. Also, psychotherapy is more successful in a relatively relaxed than in an excessively tense individual. Hence relaxation may be considered a valuable adjunct to psychotherapy.

Bibliography

1. Hadley, J. M. Various Roles of Relaxation in Psychotherapeutics. *J. Gen. Psychol.* 1938, 19, 191-203.
2. Pascal, G. R. The Use of Relaxation in Short-Term Psychotherapy. *J. Abnorm. Soc. Psychol.* 1947, 42, 226-242.
3. Schultz, J. H. *Das Autogene Training.* (6th Edition) Leipzig, Germany: George Thieme, 1950.
4. Yates, Dorothy H. Relaxation in Psychotherapy. *J. Gen. Psychol.* 1946, 34, 213-238.
5. Bagchi, B. K. Mental Hygiene and the Hindu Doctrine of Relaxation. *Ment. Hyg.* 1936, 20, 424-441.
6. Behrend, H. J. and J. Weiss. Therapeutic Relaxation. *New York State J., Med.* 1941, 41:2, 1838-1842.
7. Bouman, H. D., D. I. Briggs, and S. E. Harris. The Value of Relaxation Procedures in the Treatment of Anxiety Tension States. *Occ. Ther. Rehab.* 1950, 29:6, 345-353.
8. Jacobson, E. *Progressive Relaxation* (3rd Edition). University of Chicago Press, Chicago: 1944.
9. Hoskisson, J. B. The Theory and Practice of Self-induced Deep Relaxation. *Brit. J. of Phys. Med.* 1950, 13:4, 92-94.
10. Schade, M. L. *Relaxation.* Unpublished Master's Thesis, University of Wisconsin, 1948.

ON THE MATHEMATICAL REPRESENTATION OF THE SYNDROME

FRANK M. DU MAS

Louisiana State University

Introduction

The syndrome is a complex, multivariate phenomenon. It may be thought of as a more or less unique configuration of symptoms. The configuration of symptoms of one syndrome may be different from that of another syndrome in two basic aspects: qualitative and quantitative. For example: Syndrome, S_1, may have the qualitative symptom categories, C_1, C_3, C_9, C_{10}, C_{15}; syndrome, S_2, may have the qualitative symptom categories, C_1, C_4, C_8, C_9, C_{17}, C_{22}. S_1 may be quantitatively different from S_2 also. For example: S_1 and S_2 both have the qualitative symptom, C_9, but S_1 may have a lower degree of the symptomatic quality than S_2. That is, S_1, may be represented by a lower point on a continuum of magnitudes than S_2. Syndromes, then, may differ in the categories that are exhibited and the magnitude of categorical characteristics.

Patients exhibiting the same syndrome may be differentiated by the gross verbal numbers: mild, moderate, severe. For example, two individuals exhibit the same syndrome, S_1, and therefore the same categories C_1, C_3, C_9, C_{10} and C_{15} but one individual has a low degree of these attributes and is diagnosed 'mild'; the other individual has a high degree of these attributes and is diagnosed 'severe'. The preceding general statements apply to syndromes of any sort. The discussion that follows will be restricted to the psychiatric syndrome, but much that will be said will apply to other types of syndromes.

The typical medical or psychiatric description of a syndrome is usually verbal. This verbal description may be long or short, a paragraph or several pages. The typical description of a syndrome by a clinical psychologist is also mainly verbal but quantitative data in the form of test scores and pictographic representation in the form of profiles or psychographs are often included. We shall attempt to develop a rationale for the extension of quantitative descriptions of syndromes. Our attempt should not be regarded as a lack of appreciation for verbal descriptions but rather as an attempt to supplement these descriptions by quantitative or mathematical ones.

A More Conventional Approach to the Problem

For a complete and exact description of a human being in all respects a very great number of variables would have to be considered: organic, environmental, experiential and behavioral.

Definitions.

n_t : the total number of variables required to describe human beings in *all* respects.

$n_r = n_t - n_i$: where n_i is the number of irrelevant variables and n_r is the number of relevant variables required to describe *human personality.*

$n < n_r$: where n is the number of orthogonal variables required for an *adequate* description of human personality.

By 'relevant' we mean that a variable makes a contribution to the prediction of a phenomenon. We do not specify operationally how this is done because there are many techniques which can be applied to the same data for purposes of prediction. By 'irrelevant' we mean that a variable does not make a contribution to the prediction of a phenomenon. It is obvious that n_r will be fairly large. Its size would seem to obviate any practical application of the concepts presented here. Perfection is never attainable but approximations are. The last term defined, n, is

necessary if we are to attain an approximation to the best possible representation and prediction. This approximation should be adequate to our needs as clinical psychologists but it is obvious that it will never be perfect.

The n variables may be regarded as the dimensions of an n-dimensional sphere. Equation (1) is the general equation for an n-dimensional sphere.

$$X_1^2 + X_2^2 + \ldots + X_n^2 = r^2, \qquad (1)$$

where $X_1, X_2, \ldots X_n$ are values of the orthogonal relevant dimensions, r is the radius, and the origin is zero.

Let us suppose that the scores of the general population are normally distributed, written in the form of standard or Z scores, and the limits of each of the n variables are arbitrarily set at ± 10 Z. We can then conceive of an n-dimensional sphere whose origin is zero (the intersection of the mean Z scores) and whose radius is 10 Z. All points for the population should be locatable in this hypersphere.

The means of the syndromic group will probably be other than the population means (Z = 0) on each of the n dimensions. The means of the syndromic group define a point ($M_{X_1}, M_{X_2}, \ldots M_{X_n}$) in n space; the n scores for an individual define a point in n space also.

Any syndromic group will have a distribution and a standard error of its mean, σ_{M_X}, on each of the n dimensions or n subtests. Some sort of an average standard error of the means can be calculated. Since the standard errors of the means on the subtests will probably not be equal, the derivation of an average standard error of the means will yield a statistic which is at best only approximate. However, some notion of the reliability of the syndromic point may be obtained by simply summing the variances of the means for all subtests and dividing by the number of subtests, n, as

$$\sigma_p^2 = 1/n \quad (\sigma_{M_{X_1}}^2 + \sigma_{M_{X_2}}^2 + \ldots\ldots + \sigma_{M_{X_n}}^2) \qquad (2)$$

Where σ_p^2 is the variance for a point and the square root of this value is an average standard error of the n means.

Let the syndromic means define an origin for other hyperspheres and let $1\sigma_p^2$, $2\sigma_p^2$, $3\sigma_p^2$, ... be successive radii of these other hyperspheres. We have then a set of encapsulated or concentric hyperspheres and some idea of the reliability of the syndromic point is obtained.

An equation for the syndromic hyperspheres may be written

$$(X_1 - M_{X_1})^2 + (X_2 - M_{X_2})^2 + \ldots\ldots + (X_n - M_{X_n})^2 = \sigma_p^2. \qquad (3)$$

Where,

$X_1, X_2, \ldots X_n$ are values of the orthogonal relevant dimensions,

$M_{X_1}, M_{X_2}, \ldots M_{X_n}$ are the means of a syndromic population,

Origin: ($M_{X_1}, M_{X_2}, \ldots M_{X_n}$),

σ_p^2: the mean standard error of the n means.

It is obvious that what has been said can be done for each of several syndromes. It also applies to the profile for an individual provided some estimate can be made of the variance of that individual. If the individual cannot be retested several times, then one approximation can be made by using the standard error of measurement for each of the subtests and appropriate substitution in formulas (2) and (3).

A New Approach to the Problem

In clinical psychology, test profiles have been widely used as diagnostic aids. When evaluating a set of magnitudes the psychologist has, in the past, often used two methods: subjective evaluation of the test scores and pictographic representation. When test scores are graphically delineated we call the graph a profile. Examples of this pictographic representation are the MMPI, Wechsler and Kuder profiles.

The profile is composed of a set of subtests or dimensions. An individual's scores on the subtests are plotted and a straight line drawn to connect adjacent score points. Figure 1 is an example of such a profile. Psychologists like profiles. Is it possible to retain the profile but at the same time give it more rational significance? The reader should look at Figure 1 again. It will be noted that:

a. There is no unique sequence or order for the subtests.
b. The distances between subtest ordinates are arbitrarily equal.
c. There is no single continuum for the abscissa.
d. The profile segments are not rational continua.

Let us reason together in the hope that we can attain a conception of the profile which permits the four statements above to be written affirmatively.

Table 1. Hypothetical Data. Complete Explanation in Text.

Subtests	Z score	r_{sp}
A	1.00	.30
B	3.30	.90
C	1.00	.10
D	3.30	.77
E	2.30	.50
F	2.00	.65
G	1.60	.25

Consider table 1 as data for a syndromic group. Figure 1 is a profile constructed from data in the first two columns. Assume that we have a perfect test, p, of the syndrome; that is, p has perfect reliability and validity. Assume that we have a set of subtests, s denoting a subtest, each of which is correlated with p as shown in table 1. Assume further that the distributions of p and s are normal, scores are standard or Z scores, each test is composed of a very large number of items, and that the correlation between s and p, r_{sp}, is linear and positive. Equation (4) is the familiar regression equation.

$$Y' = r_{XY}\frac{\sigma_Y}{\sigma_X}(X-M_X) + M_Y. \tag{4}$$

Rewriting equation (4) in our notation we get

$$Z_s = r_{sp}\frac{\sigma_s}{\sigma_p}(Z_p - M_{Z_p}) + M_{Z_s}. \tag{5}$$

Since M_{Z_p}, $M_{Z_s} = 0$ and σ_s, $\sigma_p = 1$, equation (5) reduces to

$$Z_s = Z_p \; r_{sp}, \tag{6}$$

and this is of the linear form

$$Y = mX + b. \tag{7}$$

It is obvious that Z_p is the slope of a line and the ordinate intercept is 0. Suppose we have three patients: one is very sick, categorized as 'severe' and has a score $Z_p = 4$; another is moderately sick, categorized as 'moderate' and has a score, $Z_p = 0$; the third is not very sick, categorized as 'mild' and has a score, $Z_p = -4$.

We construct a profile graph by arbitrarily placing the limits of the test distributions at, say, $\pm 5Z$ and label the ordinate on the left accordingly. The limits of the abscissa are 0 and 1. The subtest ordinates are drawn in according to the correlation of a subtest with the perfect test. This has been done in figure 2. The scores on all subtests made by the 'severe' patient are predicted from

$$Z_s = 4\, r_{sp}. \tag{8}$$

The subtests' scores by the 'moderate' patient are predicted from

$$Z_s = 0. \tag{9}$$

The subtests' scores by the 'mild' patient are predicted from

$$Z_s = -4\, r_{sp}. \tag{10}$$

These predicted profiles are the plotted linear functions of figure 2.

The equation for the standard error of estimating Y from X is

$$\sigma_{YX} = \sigma_Y \sqrt{1 - r^2_{XY}}. \tag{11}$$

Equation (11) written in our notation becomes

$$\sigma_{Z_s Z_p} = \sqrt{1 - r^2_{sp}}. \tag{12}$$

since $\sigma_{Z_s} = 1$. The right side of equation (12) is the coefficient of alienation.

It is obvious that one can calculate $\sigma_{Z_s Z_p}$ for each subtest and then plot a curve through these points. One could plot many curves, say, for $\pm .5\, \sigma_{Z_s Z_p}$, $\pm 1\, \sigma_{Z_s Z_p}$ $\pm 2\, \sigma_{Z_s Z_p}$, etc. We have plotted $\pm .675\, \sigma_{Z_s Z_p}$, which corresponds to a probability, P, of .50. Within the non-linear bounds of figure 2, an empirical profile should fall 50% of the time for each of the indicated patients.

The dotted line of figure 2 is the profile plotted from the data in table 1. If the linear function denoted 'severe' is regarded as a best fit for the dotted profile, then equation (8) is the equation for the dotted profile.

Another mathematical description of the dotted profile is given by an equation for a broken line as follows:

$$\left[\begin{array}{ll}
Z_s = 4\, r_{sp} + .6, & (.10 \leq r_{sp} \leq .25); \\
Z_s = -12 r_{sp} + 4.6, & (.25 \leq r_{sp} \leq .30); \\
Z_s = 6.5\, r_{sp} - .95 & (.30 \leq r_{sp} \leq .50); \\
Z_s = -2\, r_{sp} + 3.3. & (.50 \leq r_{sp} \leq .65); \\
Z_s = 10.8\, r_{sp} - 5.02 & (.65 \leq r_{sp} \leq .77); \\
Z_s = 3.3 & (.77 \leq r_{sp} \leq .90).
\end{array}\right] \tag{13}$$

Any line can be described by the equation

$$Y = a + bX + cX^2 + dX^3 + \ldots\ldots + pX^n. \tag{14}$$

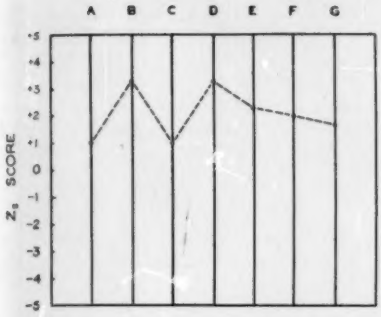

FIGURE 1. CONVENTIONAL PROFILE CONSTRUCTED FROM DATA IN TABLE 1.

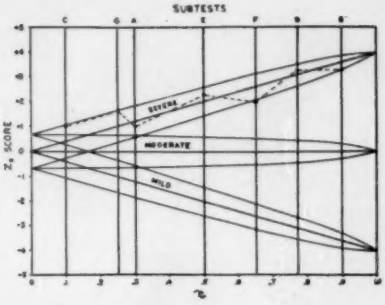

FIGURE 2. NEW METHOD OF CONSTRUCTING PROFILES. EXPLANATION IN TEXT.

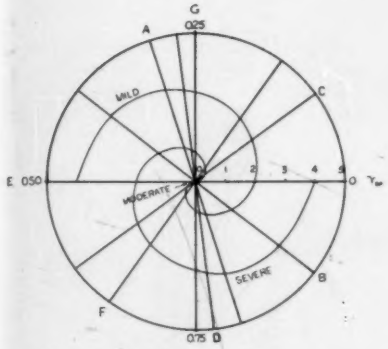

FIGURE 3. PROFILES IN MODIFIED POLAR COORDINATES. EXPLANATION IN TEXT.

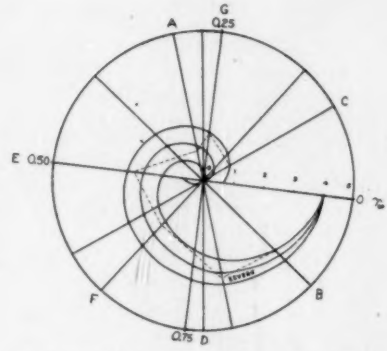

FIGURE 4. PROFILE IN MODIFIED POLAR COORDINATES. DOTTED LINE PROFILE IS THE PROFILE OF FIGURE 1. EXPLANATION IN TEXT.

Since for a given set of points an infinite number of different curves can be drawn, some criterion is needed for a 'best' curve. We suggest that such a criterion might well be: that equation of lowest degree which defines a line that passes through all the points and that satisfies the last squares criterion in some specified manner. This criterion is faulty in that it does not specify in just what way the least squares criterion will be applied. This may be decided on the basis of the degree of precision required or obtainable in practical situations.

There are other mathematical representations or descriptions of a line, but the ones given previously seem to hold the greatest promise for practical application to the phenomena of profiles. There seems to be a remote possibility that the representation and analysis of profiles in the future might be facilitated by recourse to the harmonic analyzer, an instrument used in the physical sciences to study oscillatory phenomena, or some such instrument yet to be devised. If such ever becomes possible, the machine will yield the necessary constants for the profile. Pending the advent of such a machine, an interesting problem presents itself when one considers the applicability of Fourier series and Fourier integrals to the problem of profiles. The Fourier integral approach seems to hold considerable promise.

Profiles in Polar Coordinates. The previous discussion has been restricted to profiles depicted in rectangular coordinates. An interesting innovation occurs when profiles are represented in the polar system.

The polar system requires that one choose an *origin* or *pole* and then pass a fixed line through it. This fixed line is the *polar axis* or *initial line.* A point, P, has a *radius vector,* rho, and a *vectorial angle,* Θ, as its coordinates. There are 360° around the origin. One can translate theta into r_{sp} from

$$\Theta = 3.6\, r_{sp}. \tag{15}$$

The radius vector, rho, is defined by the subtest scores. Any point, P, is defined then not by (rho, Θ) but by (Z_s, r_{sp}). Figure 3 shows the three patients' profiles plotted in this modified polar system. These curves are the old spiral Archimedes; the polar equation for such curves is

$$\text{rho} = a\Theta. \tag{16}$$

Our modification of the polar system permits us to rewrite the polar equation (16) as

$$Z_s = a\, r_{sp}, \tag{17}$$

where a is the slope in rectangular coordinates.

The modified polar profiles were constructed as follows: The pole is the mean Z_s score, that is, zero. The maximum radius vector is equal to the subtest limits, that is, 5 Z_s. The angular quantity, Θ, is translated into r_{sp}, that is, the values of the correlation are laid off around the circumference of the circle delineated by the maximum radius vector sweeping around the pole.

Figure 4 shows the 'severe' predicted profile as the midmost solid line. The outer solid lines are the bounds within which an 'empirical' profile for an individual having $Z_p = 4$ should fall 50% of the time. The dotted line is the profile of Figures 1 and 2.

Discussion

Many questions have arisen in the reader's thoughts regarding the possibility of practical application of the ideas presented here. No definite answer will be given to many of these questions but some indication will be given to show that the ideas presented in this article may have empirical relevance.

Normality of the test distribution. Any continuous distribution can be transformed into a normal distribution and meaningful Z scores can be derived for any normal distribution.

Number of items in a test. This offers no special problem. A practical definition of "a large number of items" might be, say, that the number of items in a subtest be at least 50 or 60.

Linear and positive correlation. If correlation is not linear a transformation can always be found which will make it so. All negative correlations can be made positive by simply reversing the subtest scales.

Profile segments as Continua. When subtests have a very large number of items, intermixing of the items would yield a very, very large number of hybrid tests, each different from the other. If an individual got 80% of the items on one test "correct" and 20% on another test, then if a hybrid test composed of an equal number of randomly selected items from the two original tests were constructed, the percentage of items gotten correct by the individual on the hybrid test would tend to be lower than 80%, but higher than 20%. The functional relationship of the hybrid tests scores with r_{sp} would depend on many things: scoring or weighting method, types of original tests, etc. However, since the number of different hybrid tests would be extraordinarily large, it seems that theoretically one could select a set of hybrid

tests such that the hybrid test scores were a linear function defined by the two score points of the original tests. If the number of hybrid tests was fairly large, then a profile segment could be regarded as a continuum.

Since the two original tests would be considerably different, the above rationale applied to all profile segments would lead to what the present writer [1] has called a *Catinuum*; an equation for the profile would also be an equation for the profile as a catinuum.

Correlations or Factor Loadings? The abscissa of the profile graph might be in terms of a factor, in which case the limits could be 0 and 1, and a subtest ordinate would be drawn perpendicular to the abscissa at the point indicating the factor loading of the subtest. It seems that the profile would always have positive slope when the factor loadings of the subtests were all positive. The factor might take the place of the perfect criterion test and the factor loadings take the place of the correlations, r_{tp}. The general nature of the profile curves has not been worked out, but it seems that this problem is capable of a rational solution.

Profiles in Polar Coordinates. Rectangular coordinates may not always be the clearest way to discriminate between different profiles. Also, certain problems can be handled in the form of polar equation more easily than in other forms. Since the present writer cannot foresee all the possible implications of either system, it seems that the polar system may suggest other problems and solutions that the writer does not perceive.

Perfect test and profiles. The question might be raised as to the utility of a profile composed of subtests when a single perfect, near-perfect, or very good test is available which yields a single score for each individual. We present the following arguments in favor of the profile even when a single good estimate is obtainable. First, there is no reason why the profile cannot be used to supplement a single good estimate. Second, there are many tests composed of subtests which yield a single total score and also a profile for the subtests. Psychologists tend to prefer the profile in preference to the single score. The Wechsler IQ and profile are examples. Third, since in empirical work perfection will never be obtained, a single test with high reliability and validity might well be considered as our previously mentioned 'perfect test' and a profile constructed from a set of correlated subtests used to give a new estimate of the reliability and validity of the 'perfect test'. Fourth, a different kind of clinical impression is obtained from a profile. The single score does not give the same impression. Fifth, a profile or line permits quite different mathematical analyses to be made. A point (a single score) has certain serious mathematical limitations. Sixth, the present interest in psychology is in clusters or configurations rather than single measurements. The syndrome itself, of major importance in clinical psychology, is generally conceived of as a configuration of attributes rather than a single dimension along which individuals can be ordered. The 'single dimension' concept may be reinterpreted as a unidimensional continuum which is, however, composed of or generated from several other dimensions. There is no reason why one cannot choose to use either or both the unidimensional multivariate continuum as well as the separate dimensions.

The perfect test. The perfect test will, of course, never be empirically obtainable. We can suggest several ways of dealing with the problem. First, it may be possible to construct a highly valid and reliable test which can then be regarded as an approximation to the perfect criterion. Second, a total score may be obtained from the subtests and used as the perfect test. Third, some type of factor may be located which indicates the saturation of a test with the factor. The perfect test would result from constructing a test which is highly saturated with the factor. Fourth, a unidimensional multivariate continuum could be generated from the subtests themselves. This is mathematically more precise than many of the present methods of obtaining a total score by arbitrarily weighting subtest scores. Fifth, the perfect test may be a

subjective continuum. For example, a judge or group of judges are asked to use the normal distribution as a frame of reference and estimate a Z_p score for each individual. Sixth, a measure which is or is thought to be very highly correlated with the attribute under consideration might be used as the perfect test.

This discussion is highly speculative. It should be regarded only as a source of ideas. The reader should be extremely cautious in attempting to apply these concepts until he has traced out the statistical, mathematical and clinical implications of our notions.

Summary

The syndrome is of great importance to clinical psychology. The syndrome is a complex, multivariate phenomenon. Syndromes differ from one another qualitatively and quantitatively. Syndromes may be analyzed subjectively or objectively. Mathematical representation and analysis of the syndrome is *one* approach to syndromic problems. Because of the power and simplicity of mathematical language and mathematical models, an attempt is made to talk about syndromes in this way.

A more conventional approach is to conceive of the uncorrelated dimensions of personality as the dimensions of an n-dimensional sphere. An individual or a syndromic group is represented as a point in this hypersphere. Equations for the population hypersphere and a hypersphere for a syndromic group or an individual were given.

Psychologists like profiles. For this reason an attempt is made to rationalize: (a) a unique sequence for the subtests making up a profile, (b) unique distances between adjacent subtests in the ordered sequence, (c) a single continuum for the abscissa of a profile and (d) a continuum for the profile segments. In order to do this a set of assumptions are made which also permit the profile to be described mathematically. Several equations for a profile in cartesian coordinates are suggested: A linear function, a broken line equation, that equation of lowest degree which defines a line that passes through all points of the profile and which also satisfies the least square criterion in some specified manner, and the possibility of describing and analyzing the profile in terms of Fourier series or Fourier integrals. It is suggested that profiles might be represented in modified polar coordinates. A polar equation is written for certain profiles.

References

1. Du Mas, F. M. Catinua, Catemensions, Catescales. *J. clin. Psychol.*, 1951, 7, 112-117.

THE APPLICATION OF AN EXPERIMENTAL "MEMORY BATTERY" IN THE PSYCHIATRIC CLINIC

JOHN R. KNOTT, JOHN P. UMBERGER AND INA McD. BILODEAU

Iowa State Psychopathic Hospital, University of Iowa, and Department of Psychiatry, State University of Iowa College of Medicine

Introduction

This study is concerned with the exploration of a brief battery of quantitative tests to supplement the usual clinical evaluation of "memory", and which could be used in routine examination of the patient upon his admission to the hospital. Ideally, such a device should be relatively free of regional and educational factors and should not contain items in common with other tests[10, 12] which might be required later as a further measure of memory function.

The criteria of goodness of a memory battery are perhaps somewhat difficult to define, as memory itself does not have a unit meaning. However, such a battery should discriminate between groups of persons with and without "memory defects" in the clinical sense, and should differentiate between clinical groups which would, in the psychopathological sense, be expected to show differences in invasion of the memory sphere. If sufficiently sharp, it should detect smaller areas or amounts of defect, perhaps not readily discernable at a gross clinical level, but nonetheless compatible with the total clinical picture.

For purposes of tentative definition, memory is here looked upon as *the ability to respond at a particular point in time, on the basis of, or with reference to, stimuli which were present at a prior point in time.* In order to test memory as thusly defined, it would be possible to use any of the classical methods of recognition or reproduction.

Method

The original exploratory battery used in this study was composed of seven short tests, five of which are in current clinical usage, either as individual tests or as parts of other psychological tests. These were, in the order in which they were administered:

1. *Pictures A-1* (Recall of pictured objects). This test is composed of drawings of ten common objects from the *Mental Examiner's Handbook*[13]. The subjects were shown the pictures (objects common to men[p. 156] were presented to male subjects and those more known to women[p. 158] were shown to the female subjects) and were instructed to name them. Immediately thereafter, the pictures were withdrawn and the subjects were requested (without prior warning) to recall as many as possible. This probably tests what is sometimes called "incidental memory."

2. *The Navy-Northwestern Successive Matrices.* This test was developed at Northwestern University by Hunt and French, who very kindly made it available for this exploratory study[3]. Each matrix is composed of a series of discretely presented geometric stimuli which, if superimposed, would form a design. Three parts of the design are presented in rapid succession and the subject is then required to select, on a multiple choice basis, the fourth and correct one to complete the design. Fourteen such successive matrices form the test. Satisfactory performance on the series requires the adoption and maintenance of a "learning set", as well as the ability to "remember" the parts of the particular matrix.

3. *Symbol Recall.* This test, designed by the authors, was incorporated in the battery as a visuo-motor equivalent of the auditory-vocal *Digits Forward Test.* Symbols from the Wechsler-Bellevue *Digits Symbol Test*[9] were formed into groups of from three to nine, by combining those symbols corresponding to the numbers appearing in the standard Wechsler *Digits Forward Test.* Each series, printed on its

own card, was presented for five seconds; after each presentation the subject was required to draw the symbols in the correct sequence.

4. *Pictures B* (recall of pictured objects). A second set of ten common objects[13] was presented to the subjects who were simply instructed to name them. The pictures were withdrawn and the subjects were then directed to name them from memory. In this instance, however, the subjects had reason to believe, on the initial exposure, that recall would be demanded.

5. *Digits Forward.* This test was adapted from the Wechsler-Bellevue Scale, Form I[9] and administered according to those directions.

6. *Pictures A-2.* At this point, the subjects were asked, without prior warning, to name as many of the pictured objects originally presented in *Pictures A-1* as was possible. Since the similar test, *Pictures B* had been interpolated between pictures A-1 and A-2, it was hoped that retroactive inhibition would operate to pick out those with greater susceptibility to the interpolation of similar material, i.e., those with poorer memory.

7. *The Benton Visual Retention Test* (*Form A*). This is a standard test[1] involving memory for geometric designs. The subjects were presented a printed design for ten seconds, the card was withdrawn, and they were then required to reproduce the design (s) with paper and pencil. Seven designs of increasing complexity form the test.

In each of the subtests, the score was the number of items to which correct responses were made.

During one phase of the investigation, the *Wechsler Memory Scale* [10] was administered concomitantly with the "Memory Battery". A brief test of intelligence level[7] was also administered, to all subjects.

Subjects. Three major groups of subjects were used:

Group 1 included 102 routine admissions to the Iowa Psychopathic Hospital during the fall and winter of 1948-49, regardless of diagnosis; the sole criterion for inclusion was that the patient be sufficiently cooperative to be examined. The age of the patients ranged from seven to seventy-nine years and formed a rectangular distribution.

Group 2 consisted of 50 freshman student nurses, all of whom were volunteers from a class of 107. The age range was from eighteen to twenty-three years. This group was randomly divided into twenty-five who were administered the memory battery as described above, and twenty-five who were administered the battery without *Pictures B* (in order to estimate the normal decrement on *Pictures A-2* by virtue of the inclusion of *Pictures B*). This was not a "control group", the subjects being young, bright-average to superior adults. On *a priori* grounds, they should be expected to exceed the patients on all tests.

Group 3 was made up of 48 patients, similar to Group 1, admitted to the hospital during the spring of 1949. In addition to the experimental memory battery, they were administered the *Wechsler Memory Scale* immediately after the experimental memory battery, so that the latter might be compared with a standard "memory "test.

Results

Table 1 presents the mean and standard deviation for each of the subtests, for both the young normals (Group 2) and the patients in Group 1. In all subtests except *Digits Forward*, the normals exceed the patient group at or above the 5% level of confidence, as would be expected. The failure of *Digits* to discriminate the selected bright normal group from the patients is of interest, in light of its widespread, though non-standardized, use as a test of "memory" or "attention".

TABLE 1. MEMORY SUBTEST MEANS AND STANDARD DEVIATIONS FOR CONTROL AND EXPERIMENTAL GROUPS.

Tests		Nurses			Patients (N = 102)
		Total Group N = 50	Group I N = 25	Group II N = 25	
Pictures A	Mean	7.7		7.5	5.9
	S.D.	1.1		1.1	1.8
Matrix	Mean	11.2			5.1
	S.D.	2.2			3.9
Symbol	Mean	4.8			3.8
	S.D.	1.3			1.6
Pictures B	Mean		7.8		5.8
	S.D.		1.1		1.8
Digits	Mean	6.7			5.9
	S.D.	1.3			1.4
Pictures A-2	Mean		5.0	7.32	2.5
	S.D.		1.3	1.1	2.1
Visual Retention	Mean	6.4			4.5
	S.D.	0.6			1.6

When intercorrelations were run between subjects for the normal group, few were significant even at the 5% level of confidence; those with the highest level of confidence were the ones having the greatest *a priori* expectancies, namely: *Pictures A-1* and *A-2* (without interpolation of *Pictures B*), + .64; and *Digits* and *Symbols*, + .44. The intercorrelations for the patient group were rather uniformly significant at a higher confidence level than those for the normals, probably as the result of the greater range of scores presented by this diverse group.

Of particular interest in a battery of this type are correlations with chronological age and with mental level. Two of the subtests show significant (one per cent level) negative correlations with chronological age and significant (one per cent level) positive correlation with mental level (as defined by a previously described battery). These are the *Navy Northwestern Successive Matrices* ($rho_{ca} = -.40$; $rho_{ml} = +.53$) and the *Benton Visual Retention tests* ($rho_{ca} = -.40$; $rho_{ml} = +.57$).

While nearly all of the tests used are inter-related, whether there is only one factor present in the statistical sense, is not at present known.

Since one criterion of goodness of a memory test is the ability to discriminate those diagnostic groups with memory deficit from those without, the patient group was subdivided diagnostically into (1) *psychoneurosis*, (2) *involutional and depressive psychosis*, (3) *schizophrenia*, (4) *organic psychosis*, (5) *psychopathic personality* and *primary behavior disorder*. Patients not falling into one of these categories were omitted from this present analysis. The organic patients were considered to be the criterion group. Table 2 presents the mean and standard deviation of the normals, and of each diagnostic group in the order mentioned above, for "mental level" as well as each of the memory subtests.

TABLE 2. MEANS AND STANDARD DEVIATIONS FOR MEMORY BATTERY SUBTESTS IN RELATION TO DIAGNOSTIC GROUPING

Subtests	Normal N=50	Psychoneurosis N=14	Depression N=35	Schizophrenia N=14	Organic Psychosis N=9	Psychopathic Personality N=10
Pictures—A	7.7±1.1	6.1±1.1	6.2±1.7	5.8±2.2	3.8+1.9	6.0+1.3
Matrix	11.2±2.3	5.0±3.8	4.5±3.3	7.0±4.1	1.9+1.0	6.7+4.2
Symbol	4.8±1.3	4.0±1.4	3.9±1.6	3.9±1.7	2.4+1.8	3.8+0.6
Pictures—B	7.8±1.1	6.5±1.5	5.8±1.6	5.8±1.8	3.9+2.1	6.2+1.6
Digits	6.7±1.3	5.9±1.3	6.0±1.2	5.7±2.0	5.6+1.6	6.2+1.5
Pictures—A-2	5.0±1.3* (N=25)	2.6±1.3	2.4±2.0	3.1±2.7	1.1+1.2	2.4+2.4
Visual Retention	6.4±0.4	5.3±1.3	4.3±1.6	5.0±1.7	2.9+1.1	5.1+0.8
Mental Level*	14.4±0.7	11.9±1.5 (N=13)*	12.3±1.7 (N=12)*	12.7±1.4 (N=12)*	10.9+1.7 (N=8)*	12.4+1.4

*Where *N* differs from that shown at top of column, entry made below involved mean value.

The F-test (*between psychoses—within psychoses*), using the five diagnostic groups, indicated differences significant at the 5% level for *Pictures A-1*, *Matrices*, and *Pictures B*, and, at the one per cent level for *Visual Retention*.

The differences between diagnostic groups were tested by "*t*" on these four subtests and are presented in Table 3. It may be seen, by inspection of the figures in the preceding table, as well as by these data, that: (1) patients with organic signs were significantly below other groups at the 5 per cent level of confidence, or greater, on all four subtests; (2) patients with signs of depression were, at the 5 per cent level of confidence, significantly below those with schizophrenia on *Matrices*, and significantly below patients with psychoneuroses on *Visual Retention*; (3) *Mental level*, *Symbols*, *Digits*, and *Pictures A-2* did not differentiate between the diagnostic groups; *Digits* failed to discriminate the patient groups from the superior normal group, while *Symbols* did not differentiate the normals from the psychoneurotic or schizophrenic groups.

TABLE 3. "*t*"-RATIOS OF DIFFERENCES BETWEEN DIAGNOSTIC GROUPS ON SUBTESTS PICTURES A-1, MATRICES, PICTURES B, AND VISUAL RETENTION

Differences	Pic A-1	Matrices	Pictures B	Visual Retention
*t*Psychoneurotic-Organic	3.033**	2.016*	3.486**	3.735**
*t*Schizophrenic-Organic	2.662**	3.313**	2.537*	3.284**
*t*Depressive-Organic	3.664**	1.905*	2.870**	2.526*
*t*Psychopathic personality-Organic	2.734**	2.900**	2.868**	3.198**
*t*Schizophrenic-Organic		2.225*		
*t*Psychoneurotic-depressive				2.061*

* = 5% level of confidence
** = 1% level of confidence

Pictures A-2, a test which it was hoped would be particularly useful and which was not, deserves some separate mention. It is to be noted that, in the normal sample, the group with *Pictures B* interpolated showed a significant drop on *A-2* while the group without *B* interpolated did not (Table 1). Thus, interpolation of *Pictures B* leads to greater forgetting of, or "poorer memory" for, *A*. In the patient group, the decrease from *A-1* to *A-2* was also significant. However, the difference between the decrements of the young normals and of the patients was not statistically significant. It is now thought that *Pictures A-2* was too difficult a test for all of the patients. What would happen on *A-2* if practice on *A-1* were increased beyond one trial (for example, if practice were given for three trials or perhaps to a criterion of one perfect recall) would seem to be worth investigating. Differential susceptibility to retroactive inhibition in the organic group might then be sufficiently spread to be demonstrable statistically.

With regard to the criterion of goodness relative to prediction of memory deficit in individual patients, preliminary evaluation has been made by reviewing the mental status examination in each patient's history and by an attempt to rate memory, as clinically judged by the patient's physician, as being "intact", as "showing deficit", or as "showing possible deficit." This was a difficult task, since the qualifying statements made with reference to memory might have meant that there was no particular involvement, or that there was reservation in the examiner's mind. An initially high criterion for "no deficit" (that the report could not include qualifications) was used. Any qualifications then placed the patient in the doubtful or "possible" category.

Out of the entire group of patients, seventy-three were said clinically to have intact memory. By using a cutting score of one point below the patient average on three or more of the subtests which discriminated organic from non-organic patients, (i.e., *Pictures A-1, Matrices, Pictures B,* and *Visual Retention,*) sixty-four were correctly assigned by the battery. Nine of the seventy-three patients would be suspect of memory loss by this quantitative criterion. These may be called "false positives."

Twelve of the patients had an unmistakable memory loss from the standpoint of the clinician. Of these, nine were correctly assigned by the three-out-of-four criterion. Of the three incorrectly assigned (the "false negatives"), one patient was below on two tests, one was below average on one test, and the other was above average on all four tests. This last patient had involutional melancholia, was agitated at the time of the clinical examination, but cooperated very well at the time the quantitative psychological battery was administered; hence the correlation of the original impression and the test data delt with two different conditions of behavior, and may be considered invalidated.

Sixteen of the patients were judged, by qualifying statements made by the physician, to have some possible invasion of memory. Of these, nine appeared to be below average on three of the four critical subtests. Seven were "false negatives" which were again reviewed, and in four cases it was found that the clinical qualification was: "memory seems fair". When these were deducted from the total of questionable, and definite but slight, losses, the correct assignments were made in nine out of twelve cases. Thus it appears that, when no memory deficit was found clinically, there was 87% agreement with the four subtest criterion. When a positive memory deficit was clinically described, there was 75% agreement. It might be added, parenthetically, that there is probably no more reason to regard these agreements as validating the quantitative findings than to regard the clinical data as being validated by agreement with the test results.

An additional group of patients was studied (Group 3) for comparative and validating purposes, using the experimental and the *Wechsler Memory Scale.* The mean Wechsler M. Q. was 99.8 $\pm$15.9.

The correlations between the *Wechsler Memory Scale* and the four critical subtests (*Pictures A-1, Matrices, Pictures B, Visual Retention*) were +.72 between sum of raw scores of the W. M. S.; and +.78 when *Digits* was removed from the W. M. S. Since the "memory quotient" derived from the Wechsler raw scores seems to be a non-linear transformation it was not used in the correlations, the raw score being used.

The data indicate a reliable degree of relationship between the two instruments. Most of this relationship seems to be accounted for by three of the four critical subtests: *Matrices, Pictures B,* and *Visual Retention.* With the exception of *Digits,* all correlations between the subtests and the Wechsler Scale were significant at the one per cent level of confidence.

Summary and Conclusions

This paper reports the exploration of a brief battery of seven short "memory" tests devised for use in the routine examination of mental hospital patients. Results are presented from a control group of 50 student nurses and from two experimental groups consisting of 102 and 48 routine admissions to the Iowa Psychopathic Hospital.

(1) *Pictures A-1* and *B,* the *Matrices,* and the *Visual Retention* test, discriminated a criterion group (organic) from other diagnostic groups. Conversely, *Mental Level, Symbol Recall, Digits Forward,* and *Pictures A-2* did not so discriminate. *Pictures A-2* might be made more selective by changing the method of administration.

(2) *Matrices* and *Visual Retention* both correlate relatively highly with *Mental Level,* which is itself not discriminatory, although they themselves are. This would suggest that these two tests may be related to "intelligence" and "memory" as differentiated processes.

(3) *Matrices*, *Visual Retention*, and *Picture Recall* may collectively serve to screen patients with clinically-defined memory deficits. At present the number of "false positives" seems relatively lower than the number of "false negatives", but agreement is, at the lowest, about 75 per cent between clinical assessment and test data.

(4) These four tests may be useful in forming a nuclear group for the development of a memory battery. It should be pointed out that these tests are all visual (as opposed to items in the Wechsler Memory Scale which are, with one exception, all verbal). The correlation between the experimental battery and the Wechsler Memory Scale, nevertheless, was quite high: $r = +.72$ to $+.78$, suggesting that, clinically, "memory" is not particularly related to stimulus modality and is probably a discrete process.

References

1. Benton, A. L. *A Visual Retention Test For Clinical Use.* New York: Psychological Corporation, New York, 1946.
2. Feldman, F. and Cameron, D. E. The Measurement of Remembering. *Am. J. Psychiat.* 1944, *100*, 788-791.
3. French, E. G. and Hunt, W. A. The Navy Northwestern Successive Matrices Test. *Am. Psychol.*, 1949, *4*, 268-269.
4. Graham, F. K. and Kendall, B. S. Performance of Brain-Damaged Cases on A Memory for Designs Test. *J. Abn. Soc. Psychol.*, 1946, *41*, 303-314.
5. Graham, F. H. and Kendall, B. S. Memory For Designs Test. *J. Consult. Psychol.* 1947, *11*, 340.
6. Johnson, H. K. and Lassner, R. Organic Deterioration. The Use of the Worcester-Wells Memory Test for Differential Diagnosis. *Psychiat. Quart.*, 1945, *19*, 283-194.
7. Knott, J. R., Cannicott, R. G., Umberger, J. P., and Bilodeau, I. McD. Brief Tests of Intelligence in the Psychiatric Clinic. *J. Clin. Psychol.*, 1951, *7*, 123-126.
8. Stone, C. P., Girdner, J., Albrecht, R. An Alternate Form of the Wechsler Memory Scale. *J. Psychol.*, 1946, *22*, 199-206.
9. Wechsler, David. The Measurement of Adult Intelligence. Baltimore: *Williams and Wilkins*, Baltimore, 1944.
10. Wechsler, David. A Standardized Memory Scale for Clinical Use. *J. Psychol.* 1945, *19*, 87-95.
11. Wells, F. L. Mental Tests in Clinical Practice. Yonkers: *World Book*, N. Y. 1927.
12. Wells, F. L. and Martin, H. A Method of Memory Examination Suitable for Psychotic cases. *Am. J. Psychiat.*, 1923, *80*, 243-257.
13. Wells, F. L. and Ruesch, Jergen. *Mental Examiners' Handbook.* New York; Psychological Corporation, 1945.

A FACTOR ANALYSIS OF THE BENDER-GESTALT TESTS OF MENTAL PATIENTS*

WILSON H. GUERTIN

Beatty Memorial Hospital, Westville, Indiana

Introduction

The Bender-Gestalt Visual Motor Test has been extensively used since its introduction in 1939 [2], but discussions of the test and manuals for interpretation are based largely upon clinical experience and depend relatively little upon published research with the instrument. This would seem to be largely a consequence of the difficulties in quantifying such a complex performance as is seen in the reproductions of the Bender-Gestalt designs.

Billingslea [3], in his very painstaking objective scoring method, demonstrates the extraordinary number of variables which are required to characterize these performances. It took him fifteen hours for each record, scoring on 137 indices which described the test performance. Billingslea recognized the need for reducing the number of indices and proceeded to do this by studying the consistency in the correlations between analogous performances on different figures to see what general principles could be substituted for the individual indices of the separate figures. Through the use of an experimental group of psychoneurotics, he was able to draw some conclusions about the performance of such subjects. This study by Billingslea does not disclose much about factors which might underlie the performance of individuals who are more disturbed than the subjects employed in his study. Actually very few studies have been reported in the literature which investigated the performance of schizophrenics on the Bender-Gestalt test [1].

It is quite apparent that the Bender-Gestalt Test records must be examined by the clinician through the application of a reduced number of useful factors. The intuitive establishment of these factors has depended upon Gestalt theory and were set forth in the original monograph [2]. More recently, Hutt [4] has done an excellent job in setting forth factors which should prove useful in a systematic interpretation. While the Gestalters are content to establish the conceptual frame-of-reference by an appeal to the general Gestalt theory, more objective clinicians would prefer to see these concepts established on the basis of empirical clusterings of variables.

Determination of the most important factors underlying Bender performances of all types of mental illness is an almost impossible task. Not only would the sampling of subjects have to be very extensive but the number of scoring variables included would make the task a tremendous one. The present investigation is in no way a complete study but is merely an application of factor analytic techniques to this general area of visuomotor performance.

Procedure

The subjects employed were 100 white patients at three mental hospitals*, including 33 females and 67 males. For statistical convenience the subjects were characterized by four classifications. The distribution of subjects by diagnostic groups were: 15 Not Psychotic, 23 Organic Brain Diseased, 29 Paranoid Schizophrenia, 30 Other Schizophrenia, and three unclassified psychosis.

The original quantification variables were selections from, and modifications of, those employed by Billingslea [3]. While each Bender production was examined with respect to the 60 variables initially employed, it was apparent that the frequency of occurrence for many of these was so low that they could not be studied

*The author wishes to acknowledge the cooperation of the Superintendents and staff at Beatty Memorial Hospital, Kalamazoo State Hospital and Indiana Hospital for Insane Criminals in making the data available.

TABLE 1. OBLIQUE FACTOR MATRIX*

Variables	Figures	Factors						
		Propensity for Curvilinear Movement	Poor Reality Contact	Careless Execution	Constriction	Poor Spatial Contiguity	Unidentified	Σa_{ij}**
Distortion of No. of waves	6	74	42	28	26	29	19	68
Tail arc poor	4	69	30	43	00	19	19	52
Non-closure	a, 4, 8	69	53	22	15	03	13	69
Handle angle poor	5	69	27	47	30	26	20	52
Part rotation	a, 4	65	27	63,	17	42	15	51
Loss of angles	3	62	06	35	07	16	22	47
Number distortion	1, 2, 3, 5	51	43	22	27	18	-24	46
Waves flattened	6	50	18	05	-05	13	20	39
Crossing point wrong	6	44	03	-03	-16	21	25	42
Tailing distortion	4	42	26	-23	-07	16	14	68
Columnar spacing irregular	2	41	10	22	03	18	17	19
Circularity poor	a	38	21	19	06	-07	06	22
Rotation $\geq 90°$		49	83	23	30	40	08	74
Figure order poor		47	78	28	27	38	17	64
Circle-period substitution	1, 2, 3, 5	34	70	33	26	32	16	53
Line quality irregularity		14	70	09	20	01	-29	67
Vertical slope irregularity	2	21	-37	17	29	18	09	62
Horizontal slope irregularity	1, 2	19	34	28	-09	28	04	28
No compulsive restroking		44	06	75	35	44	26	66
Linear regularity poor	7, 8	46	28	68	36	46	08	50
Vertical slope wrong	2	44	28	64	19	11	02	53
Pointed loops	6	44	53	60	18	25	-21	71
Squareness poor	a	35	-03	44	09	40	-13	42
Major arc radius wrong	4	21	-03	32	24	00	05	17
Erasure present		-12	07	-30	-11	-11	13	16
Undersize figures.		24	13	50	86	12	-36	88
$\geq$ 1 page		25	22	35	85	22	-03	80
No heavy lines		00	50	14	60	00	-07	58
Oversize figures		13	-10	-01	-60	25	-21	75
<1 page		-04	-19	08	55	-14	03	53
Lip approximation	5	31	21	27	34	04	-09	20
Handle attachment wrong	5	-06	07	00	33	03	-20	15
Disproportion of elements	a, 4	-05	14	18	-26	18	03	36
Angle between elements poor	7	35	24	45	09	82	08	69
Overlap of elements	a, 4, 8	30	29	31	03	78	20	62
Juxtaposition poor	7	34	26	35	17	68	04	47
Crossing angle poor	6	37	33	21	30	38	28	37
Fig. extend to paper edge		30	20	20	-26	18	87	81
Not oversize	1, 2	-24	-20	14	06	01	72	91
Rotation <90°		38	11	05	06	-02	67	67
Page balance poor		00	01	-02	11	-07	35	20

$\Sigma a_{ij}^2 = 21.37$

*Decimal points omitted.

**Based upon orthogonal matrix not reported here.

statistically, and so they were dropped from the study. Only those items which appeared in more than 19 per cent of the cases in the tail of the distribution were retained. Six of the variables were dropped because of obvious experimental dependence. The final 41 variables employed bear some resemblance to the original Billingslea indices. For editorial convenience these items will not be defined in detail and for those terms which seem vague or ambiguous, reference can be made to the word usage in Billingslea's monograph [3], except as otherwise indicated. To summarize, three general principles were operating in the retention of variables to be studied. They were: ease of scoring, lack of experimental dependence on other variables in the scoring, and frequent enough occurrence in this sample of subjects. An additional three items met these criteria satisfactorily but had such low communalities that they were dropped in the final analysis.

It was hoped that factors would be revealed which would be dependent upon gross inaccuracies in the records since the clinician seldom has the time to quantify his records carefully. Furthermore, it was felt that the productions of subjects, as disturbed as those employed in this study, disclosed sufficiently gross distortions so that more exact metric measures were not exactly necessary.

The data sheets on each subject were transferred to McBee punchcards for ease of tabulation. Tetrachoric correlation coefficients were graphically estimated from Thurstone's tables. The multiple-group centroid method of factor analysis [5] was employed because it is time-saving and likely to produce empirically meaningful factors. The clusters which determine these multiple-group factors were identified by a conventional method which produces fairly tight and relatively independent clusters. Estimates of communalities were obtained by using the highest coefficient in each column. The communality accounted for by each of the variables was determined from the orthogonal matrix obtained by rotating the oblique factor matrix herein reported.

Results

The cluster analysis was somewhat discouraging at first because of the difficulty in identifying tight clusters which were relatively independent of one another. This seems to be a consequence of the absence of high coefficients in spite of the fairly large number of moderately high coefficients which were present in the intercorrelation matrix. Nevertheless, six clusters were identified and six oblique factors were extracted simultaneously. This resulting six-factor oblique matrix for symptoms is shown in Table 1.

The total variance accounted for by the six orthogonal factors constitutes 52 per cent of the total variance of the intercorrelation matrix or 93 per cent of the estimated communality for these 41 test variables. The estimated communality, by taking the highest correlations in each column, was 56 per cent of the total variance of the intercorrelation matrix.

Factor loadings for diagnostic groups are shown in Table 2. It will be noted that only four correlations greater than .30 appear in the matrix. Table 3 shows the intercorrelations between the six oblique factors (Rpq). The correlations are rather high and give some additional information about the nature of the factors.

Discussion of Results

Much of the Bender-Gestalt performance can be understood in terms of factors since the communality of the intercorrelation matrix is greater than half of the total variance of the matrix. This makes it possible to employ clusters of scoring items (i.e. factors), instead of single items, in analyzing the performance of mental patients on the Bender-Gestalt.

The location of reference axes is always equivocal. The application of a conventional cluster analysis to locate the factors as well as the correspondence these factors and general psychodiagnostic and Gestalt principles, argues for satisfactoriness of the positions of these reference axes.

TABLE 2. FACTOR LOADINGS FOR DIAGNOSTIC VARIABLES

Factor	Diagnostic group			
	Non-Psychotic	Organic B. D.	Paranoid Schizophrenia	Other Schiz.
Propensity for Curvilinear Movement	.00	−.10	−.05	.24
Poor Reality Contact	−.27	−.39	.25	.09
Careless Execution	−.07	−.16	−.09	.13
Constriction	.33	−.04	−.05	−.22
Poor Spatial Contiguity	.25	−.33	−.14	.24
Unidentified	.35	−.20	.00	.28

TABLE 3. INTERCORRELATIONS OF THE OBLIQUE FACTORS

	Factors*					
	A	B	C	D	E	F
A	1.000	.480	.650	.211	.437	.258
B	.480	1.000	.313	.342	.367	.043
C	.650	.313	1.000	.433	.490	.177
D	.211	.342	.433	1.000	.147	−.197
E	.437	.367	.490	.147	1.000	.143
F	.258	.043	.177	−.197	.143	1.000

*The capital letters symbolize the oblique factors in the order of their presentation in the text.

It should be emphasized that these factors were obtained from a disturbed group of patients. Very few of the productions were from individuals with what might be termed mild neurotic disorders. This will tend to disclose factors which depend chiefly upon such things as poor reality contact, but the sampling is such as to exclude other possible factors such as one attributable to anxiety. A factor analysis of a neurotic sample would be essential before anything like a complete conceptual scheme of test evaluation could be proposed.

These six factors have been named inductively by considering the symptoms which showed the highest factor loadings:

The *Propensity for Curvilinear Movement Factor* is characterized by distortion in figures due to a disturbance in curvilinear movement. This is reflected on the one hand by an excessive number of waves in figure 6, and on the other by rotational tendencies seen in actual rotations or more particularly in rotations of parts of the figures. Non-closure of elements is also important here and seems to be based chiefly on the fact that the non-closures in this sample were almost entirely related to difficulty with the curvilinear element of figure 4. Underlying this clustering of test variables would seem to be a poor emotional control which makes it difficult for the

individual, having perceived the Gestalt, to maintain it. The flux or instability seen where this factor is operating can be observed clinically in schizophrenics particularly of the catatonic subtype. It will be noticed that a factor loading of .24 for Other Schizophrenics was the largest loading for any of the diagnostic groups. This factor seems to be quite important as evidenced by the sizable loadings on an appreciable number of items. Bender, herself, makes much of this as is seen in her statement, "the inherent quality of all optic images is movement."[2, p. 46] This movement tendency is a strong feature of most Gestalt principles and is seen by Bender and others as being basic to all visuomotor organization. The primitive features of such movement are found genetically in the view that small circular movements are the elements upon which a child bases the later development of more complex perceptual movements. Hutt[4, p. 10] discusses modifications of curvature as a basic factor in the modifications of the Gestalt. He attributes such modifications to deviations in emotional functioning. The present findings support such a view.

The *Poor Reality Contact Factor* is one which is familiar to all who work with psychotic patients. Among the important items here are the inability to maintain sufficient reality contact to reproduce the figures in even an approximate fashion and associated features which reflect this aspect in an inability to perceive or meet the general requirements of the environment. The so-called regressive circle-period substitutions are encountered as an important feature of this factor. The Paranoid Schizophrenics show the highest positive loading for this factor as might be anticipated, while non-psychotics are negatively loaded with this factor. The highest relationship was for the Organic Brain Diseased group who were loaded −.39 on this factor. This would suggest that the Organic Brain Disease cases are capable of fairly accurate perception of the stimulus material with errors confined to smaller executional aspects of the test. It would seem worth reporting the individual correlations between diagnostic types and the rotation of the figure 90 or more degrees. These correlations give support to Hutt's views[4, p. 9] on the importance of rotations in psychoses. The correlations with Rotation are as follows: Not Psychotic, −.25; Organic Brain Disease, −.33; Paranoid Schizophrenia, .34; and Other Schizophrenia, .06. In the case of Rotation it appears that there is an element of active distortion present rather than just disorientation since the correlation is highest for the Paranoid Schizophrenics.

The *Careless Execution Factor* is characterized both by a lack of care taken in reproducing accurately and in compensating for such inaccuracies by such things as restroking. This factor would be important for careless individuals who are not ego-involved in their reproductions either because of a character disorder or psychosis. As far as is known, such a factor has not been formally proposed in spite of its obvious importance. Poor motivation or impulsiveness can exist in almost any type of case so it is not too surprising that the loadings for each of the diagnostic groups on this factor are all quite small.

The *Constriction Factor* embodies the well-affirmed psychodiagnostic principle that insecure people will constrict the magnitudes of their movements in a formal testing situation. This factor is largely composed of size variables which have varying degrees of experimental dependence. The only heavily loaded variable which is not concerned with size is the absence of heavy lines. The whole picture, then, is one of timidity with resulting constriction. So much attention has been given to this factor in the literature that it seems unnecessary to dwell further on it.

The *Poor Spatial Contiguity Factor* indicates a difficulty in juxtaposing the two elements of the figure properly. The factor is characterized nicely by Hutt as, "difficulty with crossings and closures"[4, p. 12], and is ascribed to psychoneurotics. It is seen by the diagnostic loadings that on the one hand Organic Brain Disease is negatively related to this factor, while on the other, the Non-Psychotics and Other Schizophrenics are loaded positively with this factor. The psychoanalytic interpretation

of symbolic representation of conflict when two elements are juxtaposed may have some confirmation in these findings.

The *Unidentified Factor* is composed of only three significantly loaded items. The nature of this factor is wholly unclear and the number of items loaded are relatively few. It seems desirable to disregard this factor in the present study and await further data before trying to understand it.

Five group factors seem to be fairly well established by the present study but probably represent a relatively small portion of the total common factor space that could be identified. A better picture of this factor space could be obtained by studying all aspects of test performance for all types of subjects. Breaking the general area of diagnosis down into separate factor analyses on samples of organics, schizophrenics and psychoneurotics would seem to be desirable. Nevertheless, it is difficult to see how a similar extensive evaluation of the numerous performance indices can be made on the Bender-Gestalt test itself. A factor analysis of each figure would be a more thorough attempt at understanding the principles involved but would ignore interrelationships between figures. It would be necessary therefore to do an additional analysis of the more general aspects of the test in relation to the factors revealed on each of the figures. This would mean a total of 10 factor analyses, one for each figure. These 10 factor analyses should probably be repeated with the three separate diagnostic groups in order to provide a thorough study making a total of 30 factorings in all.

In conclusion, it is felt that factor analysis proves to be quite a laborious procedure for establishing a conceptual scheme for interpretation of the Bender-Gestalt but the approach has been demonstrated in principle and some contributions relative to the empirical clustering of performance variables have been made.

Summary and Conclusions

1. The quantification procedures of the Bender-Gestalt in the literature suggested the possibility of factor analyzing the objectively scored test variables in an effort to see which ones clustered together. The study is a restricted one because of the large number of possible variables which might have been employed.

2. The Bender reproductions of 100 mental hospital patients were evaluated in terms of a number of semi-objective scoring indices. The intercorrelation matrix of these variables was factored with the multiple-group centroid factor analytic technique. Factor loadings are presented for 41 test variables and four gross diagnostic categories.

3. The five oblique factors obtained which seem to have meaning were as follows: *Propensity to Curvilinear Movement, Poor Reality Contact, Careless Execution, Constriction, and Poor Spatial Contiguity.*

4. The application of factor analysis in establishing a conceptual scheme for the interpretation of Bender-Gestalt productions is illustrated and the difficulties in terms of the labor involved in a more thorough study are discussed.

Bibliography

1. Abt, L. E., and Bellak, L. *Projective Psychology.* New York: Alfred A. Knopf, 1950.
2. Bender, Lauretta. A visual motor Gestalt test and its clinical use. *Amer. Orthopsychiat. Assoc., Res. Monogr.*, 1938, *No. 3.*
3. Billingslea, F. Y., The Bender-Gestalt: An objective scoring method and validating data. *Clin. Psychol. Monogr.*, 1948, *No. 1.*
4. Hutt, M. L. *A tentative guide for the administration and interpretation of the Bender-Gestalt Test, U. S. Army:* Adjutant General's School, 1945 (Restricted).
5. Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univ. of Chicago Press, 1947.

RELATIONSHIP OF RORSCHACH SCORING CATEGORIES TO MODES OF PERCEPTION INDUCED BY INTERMITTENT PHOTIC STIMULATION—A METHODOLOGICAL STUDY OF PERCEPTION**

RUTH G. MATARAZZO ROBERT I. WATSON GEORGE A. ULETT*

Department of Neuropsychiatry
Washington University School of Medicine
St. Louis, Missouri

Introduction

Although Rorschach interpretation has long been based upon several empirical assumptions about the meaning of such scoring categories as C, F, and M, relatively little experimental work has been done in regard to discovering whether or not there are characteristic psychological or physiological correlates of these factors. Thus the Rorschach test, after more than 25 years, is still largely neglected from the standpoint of experimentally validated perceptual and personality correlates. A problem preliminary to discovering the exact personality correlates of the factors in question is to determine whether the Rorschach categories are actually meaningful to perception in general, and whether a subject's perception tends to remain typical in other situations which would elicit behavior that could be scored in terms generally comparable to those of the Rorschach situation. As Bruner[3] says, "Ink blots are no separate species of perceptual stimulus."

In a preliminary study Ulett[8] reported the various subjective sensations noticed by subjects stimulated with flickering light. His control subjects reported seeing color, movement and patterns in response to the flickering white light and a few had well-organized visual illusions. The reported patterns changed with alteration of the stimulus in regard to intensity or rate of flicker. In addition to these experiences, some psychiatric patients reported that symptoms related to their illness, such as dizziness and falling sensations, were experienced during the exposure to flickering light. It was noted that there seemed to be a frequency specificity in various patients and that when a critical rate of stimulation was reached, driving of the EEG pattern was enhanced and the subjective sensations were more marked. The relation of such phenomena to personality factors was investigated by Ulett[9] who found that it was possible to discriminate fairly well between normal and anxiety-prone individuals according to the kind of sensations experienced under photic stimulation. Anxiety-prone individuals were more likely to develop feelings of dysphoria.

Since the subjects in the above studies of intermittent photic stimulation reported their induced sensations in terms which could be scored in a way comparable in some respects to the categories scored in the Rorschach situation, and since emotional and psychiatric symptomatology were evoked by a stimulus producing simultaneously a change in their brain wave patterns, it seemed that here possibly was a way of relating determinants of perception utilized in the Rorschach test to the modes of perception in another situation in which neurophysiologic parameters were more accessible.

Subjects and Procedure

In the present study 181 male subjects were used, 149 of whom were volunteers with a minimum of eighth grade education drawn from the Air Force reserves, regular Army personnel, university and seminary students. The remaining 32 subjects

*This report was completed under AF Contract Number 33(038)-13884. It is one of a series of studies under D and R Project 21-37-002, Development of Psychiatric Screening of Flying Personnel, Department of Clinical Psychology, United States Air Force School of Aviation Medicine.

**The authors wish to express acknowledgment to Miss Ann Lawler who administered and scored all Rorschach tests, to Dr. George Winokur who did the psychiatric screening, and to Miss Rosemary Baessler who conducted the intermittent photic stimulation experiment. Thanks are also due to Dr. Goldine Gleser for statistical consulation.

were patients of similar age, sex and education selected from the hospital and clinic of the Department of Neuropsychiatry of the Washington University School of Medicine and Barnes Hospital. These latter had been diagnosed as having anxiety neurosis or some psychiatric disease in which anxiety was a major feature. All subjects were part of a larger study of anxiety-proneness[9] and hence had been carefully screened both psychologically and psychiatrically for evidence of personality disturbance. All were subjected to a three-hour battery of psychological tests including the Rorschach, and were, in addition interviewed by a psychiatrist for one hour. For both the aforementioned situations they were rated on an 8 point scale for anxiety-proneness. Ratings of 1, 2, and 3 were considered normal, 4 and 5 borderline, and 6, 7, and 8, abnormal. Where the psychologist's and psychiatrist's ratings were at variance, a conference was held and a group rating decided upon, this rating usually being an average of the two original ratings. As a result of such screening, 133 individuals were distinguished as being free from any significant anxiety (i.e., rated 5 or below). Sixteen of the "normal" volunteers were considered to be very prone to develop anxiety under stress (rated 6, 7, and 8). The third group consisted of the aforementioned 32 psychiatric patients.

The Rorschach test was administered first in all cases and, several days later, the intermittent photic stimulation experiment was carried out in the electroencephalographic recording laboratory. The method reported by Ulett[8] was used in the latter situation. The subject was seated and a flashed-opal glass screen 30″ in diameter placed about six inches from his face. Thus the flickering light filled the visual field almost entirely. Directions given to the subjects were to report "any sensations experienced other than the flashing off and on of a white light." Frequency of stimulus was varied over 22 frequencies from 2 to 30 flashes per second presented in ascending order. Each frequency was maintained for forty seconds followed by forty seconds of rest. An attempt was made to elicit some description of the subjective responses to each flicker frequency. All such sensations were recorded verbatim.

Rorschach protocols were scored according to the system of Hertz. The reports of the subjective sensations induced by photic stimulation were scored as similarly as possible to Hertz' method of scoring for the Rorschach and without knowledge of the Rorschach results. Criteria for the scoring categories used in evaluation of this material are as follows:

C was defined as any mention of color. It was felt necessary to differentiate between a response of one color and many — the latter warranting a greater color score — and so a separate C was given for each color mentioned in each response. A response of "many colors" was credited with one C. No FC or CF categories were used since, in this methodological study, it was deemed impossible to differentiate accurately among these three color categories.

CR the number of different frequencies at which there was some color response.

C' included any mention of black or white (other than on and off flash of the light), for example; shades of gray or black, black venetian blinds, black and white design, or streak of lightning.

F_1 was the scoring used for any fairly well defined object such as a cross, swords, an airplane wing, turtle with a top hat, waterfall, pages of a book, honeycomb, fan, or wheel and spokes.

F_2 signified insufficiently described forms such as designs, figures and objects.

m included movement of any object or organism, for example; moving geometrical figures or a flock of seagulls flying. No attempt was made to differentiate between inanimate movement and human or animal movement as the incidence of the latter types of response was extremely rare.

FK was the scoring used for vista or suggestion of three-dimensional expanse, for example; movement toward or away from self, a tunnel or a view from an airplane.

c was defined as surface appearance or diffusion, for example; surface of the moon, shimmer, vibration, clouds, blue haze or white blur.

TABLE 1. TESTS OF SIGNIFICANCE OF DIFFERENCES IN MEANS FOR DIFFERENT GROUPS ON THE PHOTIC VARIABLES

Flicker Variable	*t* between normal-anxious normal	*t* between normal-anxious patient	*t* between anxious normal-anxious patient
C	1.26	.44	1.40
CR	.77	.31	.51
C'	.02	.42	.29
F_1	.78	1.05	1.37
F_2	.03	1.26	1.06
m	1.63	1.05	2.08*
K	.37	.61	.71
FK	1.65	1.25	2.00
Stereotypy	.89	1.87	.53
No Response	.42	.11	.20
Unusual Response	1.26	.64	.96
	df = 147	df = 163	df = 46

* Significant at .05

The above categories and criteria were selected after preliminary scoring of the first fifty protocols from the photic stimulation experiments, and correlation of the major factors with corresponding Rorschach data. For such comparisons we used all of the Rorschach scoring categories which appeared in our protocols with sufficient frequency to warrant analysis. Although there was some difficulty in matching the two different types of scores, we initially tried all that seemed to have some similarity. Many comparisons were abandoned early, especially those flicker categories in which there were very few responses, or Rorschach categories which were too refined for the type of reports given during photic stimulation. The present criteria are still largely *a priori* in origin and could undoubtedly be more precisely related to Rorschach factors in further empirical studies. This paper can be considered only a first attempt to find the scoring categories which might be comparable to those of the Rorschach test.

RESULTS

Before ascertaining any relationship between the two sets of variables, a preliminary comparison of mean normal and anxious photic scores was made to see whether there might be a real mean difference on any factor for the groups. Mean values for the three groups (normal controls, anxiety-prone controls and patients with anxiety) were found and *t*-tests computed. Only one significant mean difference was found. The anxious patients rated significantly higher on photic "m" score than did the anxiety-prone controls. However, since this *t*-was barely significant at the .05 level of confidence, and 33 *t*-tests were made, the difference could be expected by chance. Moreover, any conclusions based on the *t*-tests should be tentative since the distribution showed marked positive skewing.

It was then necessary to determine whether those individuals who scored relatively high or low on any Rorschach variable tended to rate high or low also on the corresponding photic variable. For example, does the individual with a relatively large Sum C tend to report much color under intermittent photic stimulation? Therefore, the median of the total distribution of each Rorschach and photic variable was ascertained and every subject was rated "above" or "below" the median on each factor. In cases where the median fell within a group of scores at one point on the distribution, division was made above or below that point, depending upon whether

the median fell toward the upper or lower part of the group in question. Chi-square was then used to determine whether there was a trend toward consistency of response in the two situations. Column 1 of Table 2 shows those factors which were compared. For the total group, Rorschach Sum C and flicker C were related p = <.05. Rorschach M and flicker movement responses were related p = <.001.

TABLE 2. TESTS OF SIGNIFICANCE OF RELATIONSHIP BETWEEN RORSCHACH AND INTERMITTENT PHOTIC VARIABLES

Rorschach Variable	Flicker Variable	Chi-Square Values			
		Total N=181	Normals N=133	Anxious Patients N=32	Anxious Normals N=41
ΣC	C	4.46*	5.12*	.25	.72
ΣC	CR	2.41	2.28	.65	.00
F+	F_1	1.46	.50	1.13	1.18
M	m	18.09***	11.36***	4.09*	2.94
C'	C'	.24	.22	.62	.50
FM	m	1.21	1.13	.01	.54
F+	m	2.01	.34	2.03	1.98
FK	FK	.00	1.84	7.66**	3.96*
F+	F_2	3.32	.18	4.98*	.01

* Significant at .05 level of confidence.
** Significant at .01 level of confidence.
*** Significant at .001 level of confidence.

From the new data it appeared that the patients were contributing to some of the relationships differentially from the normals. Consequently the patients were separated and the remaining relationship for the entire normal group is summarized in column 2 of Table 2. The C and M factors were found statistically to remain about equally related to their photic counterparts.

Separate chi-squares were determined for the patient group as summarized in column 3. The normal group corresponds rather closely to the total as is to be expected since this was by far the largest sub-group included in the latter. The patients however, show some rather surprising deviations. In this group the relationship of Sum C to C apparently disappears and the correspondence between M and photic movement is attenuated (p = <.05). A relationship between Rorschach FK and photic FK becomes apparent in the patient group (p = <.01). In the patients also, the sum of Rorschach F+ responses becomes related to F_2 or poorly defined form. (Actually, F_2 probably should have a closer relationship to F+ than should F_1 since more clear-cut illusions would be the opposite of form-boundness.)

To further examine these differences between the controls and anxious patients it was decided to study those control subjects who had been classified as "anxiety-prone." As there were only 16 such subjects (a group too small for using chi-square) it seemed logical to include from the "normal control" group 25 additional subjects who had received scores of 4 on the 8 point scale for anxiety-proneness. Thus although they were grouped with the normals in a dichotomous scale they could nevertheless be considered on the upper fringe of normal as both the psychologist and psychiatrist, in their rating, had called subjects rated 4 and 5 "borderline." Column 4 of Table 2 summarizes the findings in this group. It can be seen that, as in the anxious patients, the relationship of Sum C (Rorschach) to C (sensations evoked by photic stimulation) disappears as does also the relationship between the movement categories. A relationship between Rorschach FK and photic FK is apparent but less so than in the patient group.

To find the degree of relationship between those variables which had already shown some inter-test consistency, it was decided to use some correlation technique. Extreme positive skewing, especially large on the flicker variables, fostered sus-

picion that some spurious variability at the far right might tend to have undue influence on any correlations. Consequently, a seven-point coding scale was adopted which tended to make the distributions more nearly normal. It was impossible to use standard scores for complete normalization as too many scores were piled up at the low end. Pearson product-moment correlations were used as presented in Table 3.

TABLE 3. CORRELATIONS BETWEEN CORRESPONDING RORSCHACH AND FLICKER VARIABLES. N = 181

Rorschach Variable	Flicker Variable	Pearson *r*	p	*eta*	Chi² test of linearity
ΣC	C	.17	< .05	.23	N.S.
M	m	.22	< .01	.28	N.S.
F+	F	.16	< .05		
FK	FK	.05			

This table shows three of the factors to be significantly related, although the correlations are low, ranging from .16 to .22. M and m still have the highest degree of relationship, Sum C to C a somewhat lower relationship, and here again we find Rorschach Sum F+ to be related to photic F_2. Rorschach FK shows no relationship to flicker FK for the total group. Since slight deviations from linearity had not been ruled out, eta's were determined for Sum C, M, and their corresponding photic factors. The chi-square test of linearity indicated that relationships were linear.

DISCUSSION OF RESULTS

The hypothesis being tested, namely, that the determinants of perception operative in the ambiguous Rorschach situation are related to the determinants of perception in the similarly ambiguous photic stimulus situation, appears to be borne out, at least for certain categories. Results would indicate that there is, to some extent, perceptual constancy for the individual in different situations. The unmeasurable differences in the two stimulus fields probably make the obtained similarities of response the more significant. As stated in the presentation of results, chi-square tests indicated that, in the total group, those who had a tendency to respond in terms of color or to movement categories on the Rorschach demonstrate a similar tendency in the photic stimulation experiment. It is also interesting that FK, a presumed measure of anxiety, is related in the two situations for the anxious group although not related for the "normal" controls.

The findings reported in other studies seem to substantiate the use of "M" as a factor of a constant degree of importance for the individual in many perceptual situations. For example, Klein and Schlesinger[5] found that those who scored high in Rorschach M also tended to experience apparent movement more easily in an experimental situation. Although there is considerable interest in the "C" responses the reported studies are of a clinical nature and few relate to color theory. Unfortunately, there is a paucity of experimental work in the literature for the other Rorschach variables. The lack of positive results, in the present study, for these variables indicates that perhaps the other Rorschach categories are not "pure" i.e., should be defined in a different manner, or that they represent less stable and fundamental aspects of personality. In this connection it is to be recalled that initially the ink blot test was designed as a means for studying perceptual "types" and that Hermann Rorschach was primarily concerned with the interrelationship of movement and color responses. It is to this aspect of the Rorschach test situation that the present study has suggested a possible approach.

It must, of course, be emphasized that since so little is known of what underlies the apparent predictive ability of the Rorschach, we do not possess methods of determining in an exact manner how comparable are the two methods used here.

One major difference however between the Rorschach and photic stimulation situations may be the subject's interpretation of the test. The procedure of photic

stimulation is not likely to be considered an evaluation of imagination or of any other "mental" attribute. It is more likely regarded as part of a physiological procedure and hence the subject would not be "on guard" against revealing anything which he knows or fears to be true of himself. Also, the relationship between examiner and examinee is minimal (since the subject merely reports what he sees and the examiner records) in determining the results of photic stimulation. This makes it less probable that significant motivations, defenses, interests, etc., are brought into the situation since there is little interpersonal stimulation or threat.

In this regard, Schachtell [7] states that interpretation of the Rorschach situation is an important factor in determining the subject's response. The individual's sensitivity to authority "may lead to a considerable variety of subjective definitions ... of the Rorschach situation." If, as is stated by Rorschach, among several other types of patients, the anxious are likely to take the test seriously, and, if, as we have assumed, the photic stimulation is interpreted as non-threatening by all subjects, we might expect to find a greater discrepancy between the flicker and Rorschach scores among the anxious than among normals. We find that this is true, and the distribution would possibly be even clearer were it not for the variable amount of dysphoria that may be induced during the process of photic stimulation. One very noticeable difference between normal and patient groups is that despite the significant relationship in color scores for the normals there is complete absence of relationship for color in the two anxious groups. It is plausible that "C," interpreted as spontaneous emotional reaction, should be inhibited in a situation interpreted as threatening by the subject.

It is known that intermittent photic stimulation alters the EEG [1] [8] [10]. Since it has been shown [8] that some subjects tend to report the most vivid illusions at the point where this alteration is most marked, some physiological interpretation of the present findings seems justified. It has long been known that brain tumors adjacent to the visual area, and stimulation of exposed cortex of areas 18 and 19 can produce the sensation of stars, bright colored lights, etc. [4]. Also in epileptic subjects whose abnormally discharging focus may lie in the occipital region, the aura is frequently one of visual sensations. Therefore, it is not surprising that such a stimulus as we have used, which is known to affect the visual cortex [2], would produce the visual experiences which have been described above. Grey Walter [10] has suggested that, when other than visual sensations occur with photic stimulation, the effect of stimulation is spread from the primary receptor area in the visual cortex to other regions of the brain. Therefore it seems reasonable to think that the individual's characteristic specificities of response in the two stimulus situations studied here may have their derivation from an individual patterning of neural structures.

Summary and Conclusions

The present study was undertaken to discover whether or not the Rorschach categories are meaningful to perception in general, i.e., whether an individual tends to experience another ambiguous situation similarly to that of Rorschach testing. To determine this, 181 subjects were given the Rorschach test and were later subjected to intermittent photic stimulation. The data from the second ambiguous situation were scored in a manner comparable to Rorschach categories. The conclusions derived were:

1. In the control population, the quantity of color and movement as scored on the Rorschach test responses was found to be related to the amount of color and movement produced in the subjective sensations resulting from intermitent photic stimulation.
2. The above relationship is disturbed in patients with anxiety and in control subjects who were considered anxiety-prone.
3. Among patients with anxiety and anxiety-prone control subjects, the amount of FK response on the Rorschach was found to be related to a similar perceptive reaction induced by intermittent photic stimulation.

4. The finding of a relationship between photically induced sensations and Rorschach variables may well lend some hope for the validation of a personality theory based on perceptual types as proposed from Rorschach's early experiments.

Bibliography

1. Adrian, E. D., and Matthews, D. H. The Berger Rhythm: Potential changes from the occipital lobes in man. *Brain*, 1934, 57, 355-385.
2. Bartley, S. N. *Vision, a study of its basis.* New York: D. Van Norstrand, 1941.
3. Bruner, J. S. Perceptual theory and the Rorschach Test. *J. Persona.*, 1948, 17, 157-168.
4. Fulton, J. E. *Physiology of the Nervous System.* New York: Oxford University Press, 1943.
5. Klein, C. S., and Schlesinger, E. J. Perceptual attitudes toward instability: I. Prediction of apparent movement experiences from Rorschach responses. *J. Persona.*, 1951, 19, 289-302.
6. Rorschach, H. *Psychodiagnostics.* 4th Ed. Berne, Switzerland: Verlag Hans Huber, 1949.
7. Schachtell, E. C. Subjective definitions of the Rorschach Test situation and their effect on test performance. *Psychiatry*, 1945, 8, 419-448.
8. Ulett, G. A. Preliminary observations on the use of the photically activated EEG as a research tool in psychiatry. *Psychosom. Med.*, (to appear).
9. Ulett, G. A., Gleser, G., Lawler, A., and Winokur, G. Development of an EEG index of anxiety tolerance by means of photic stimulation—its validation by psychological and psychiatric criteria. School of Aviation Medicine reports (in press).
10. Walter, W. G. The functions of electrical rhythms in the brain. *J. Ment. Sci.*, 1950, 96, 1-30.

LEVELS OF DATA IN RORSCHACH INTERPRETATION

CALDWELL, BETTYE MCD., ULETT, GEORGE A., MENSH, IVAN N., AND GRANICK, SAMUEL

Washington University School of Medicine

Introduction

Although widely accepted as a useful clinical tool, the Rorschach technique is not firmly established as a scientific method of personality assessment. Correct interpretation by means of this technique depends on the validity of many different hypotheses, all of which can ultimately be reduced to some type of numerical data, including as they do subtle observations of behavior which may contribute heavily to the accuracy of interpretation. This reminder has been frequently given clinicians in an attempt to sober the claims made for the instrument. Devotees of the method have, however, been loathe to consider it a criticism, and one has stated that in using the procedure ". . . insight and intuition must be given a free hand." (1, p. 513)

However brilliant and penetrating are interpretations made from artistic use of the method, it would seem necessary to continue attempts to remove insight and intuition as much as possible and replace them with scientific objectivity. One method of accomplishing this is to establish more firmly the relevance and usefulness of the quantitative categories into which the responses may be classified or to determine the extent to which behavioral cues contribute to personality analyses.

In an earlier publication Ulett recognized the need for such an effort and proposed the following approach:

> "The role played in interpretation by factors other than the amount and interrelationship of scoring categories has not been clearly elucidated. . . A useful approach to this problem might be to separate, in the Rorschach interpretative report, those interpretations based upon the summary of scoring data alone from the form and content of the subject's replies and to include as a separate report observations of behavior made during the test administration." (2, p. 548)

In the present study these various approaches have been conceptualized by the authors as representing three levels of data which are available to the clinician who uses the Rorschach technique. At one level there is the subject's test behavior in its complex, including the specific content of the responses and the scoring symbols

into which they may be translated. At another level, the interpreter may have access to only a transcript of the responses and must attempt to understand these without having actual face to face contact with the subject. At the final level only the scoring summary or psychogram remains, and interpretations must be made without benefit of either the behavioral data or the actual content of the responses.

To establish how much a reduction in level of information reduces the efficiency of interpretation of Rorschach data is important both for the practicing clinician and for the person interested in clarifying Rorschach methodology. Not only would this help to determine to what extent the clinician is basing his interpretation on knowledge of the correlates of certain types of gross behavior or on his general intuitive ability, but it would also shed some light on the meaningfulness of the various scoring categories and upon whether they provide efficient substitutes for the rich, qualitative material present in the subject's behavior and in the language and content of his responses.

Design

The levels to be investigated were defined as follows:

Level I. The interpreter administers the Rorschach technique to a subject, observes the subject's reaction to the procedure, makes note of any incidental verbalizations, feelings, etc.

Level II. The interpreter has access to a copy of the complete test protocol, which he may or may not then convert into scoring categories depending on his own personal preference of interpretive procedure.

Level III. The interpreter is given a copy of the scoring summary only, including the number of responses falling in different location, determinant, and content categories, average reaction time for the achromatic and chromatic cards, and number of popular and original responses.

The study was designed so as to permit measurement of the effects of the various levels in two ways. The first procedure consisted of comparing the consistency of ratings made by three different interpreters all having access to the same level of data with those made when the levels were varied. This necessitated the use of an experimental group in which levels were varied and a control group where they were held constant. Subjects for the experimental group were 24 unclassified psychiatric patients being studied in a psychiatric hospital which admits patients whose disorders range from minimal to severe. The control group consisted of 10 comparable subjects whose completely recorded protocols were drawn at random from the psychological test files. None of these 10 subjects had been examined by the three psychologists serving as judges in the present experiment. The raters were all trained psychologists with a minimum of three years of clinical experience.

Table 1 describes the program of testing of the experimental group, in which an effort was made to randomize the influence of the individual raters by having each rater see 8 subjects at Level I, 8 at Level II, and 8 at Level III.

Table 1. Outline of Testing Program of Experimental Group (N = 24) Permitting Systematic Variation of Levels of Data

	Level I	Level II	Level III
Psychologist A	1–8	9–16	17–24
Psychologist B	9–16	17–24	1–8
Psychologist C	17–24	1–8	9–16

The second procedure for determining the effects of the levels was to compare interpretations made on the same subject by raters having different levels of data with some outside criterion. That is, would an interpretation based on Level I

agree more closely with the criterion than an interpretation based upon Level III? The criterion chosen was that of psychiatric ratings made on the 24 subjects of the experimental group. In each case the psychiatric rating was made by a psychiatrist who had not only interviewed the subject for at least an hour but had also conferred with the referring doctor and the resident physician who had observed the subject's ward behavior.

Figure 1 presents an abbreviated copy of the rating form which was designed for the study and which was used with both the experimental and control groups.

I.	INTELLECTUAL FUNCTIONING						
1.	Intellectual level	−70	0	1	2	3	110+
2.	Estimate of intellectual potential	−70	0	1	2	3	110+
3.	Characteristic approach to problems and experience—quality of thought processes	hypercritical	0	1	2	3	uncritical
4.	Range of adjustive behavior	suggestible	0	1	2	3	rigid
5.	Ability to generalize	lacking	0	1	2	3	excessive
6.	Respect for reality	lacking	0	1	2	3	excessive
7.	Range of interests	stereotyped	0	1	2	3	broad
8.	Unusual material	predominant	0	1	2	3	none
II.	CONATIVE FUNCTIONING						
9.	Activity level (amount of activity)	hypoactive	0	1	2	3	hyperactive
10.	Rate of ideas (flow of speech and thought)	inhibition	0	1	2	3	flight of ideas
11	Amount of wish activity (creative imagination, fantasy)	lacking	0	1	2	3	excessive
III.	AFFECTIVE FUNCTIONING						
12.	Emotional control	none	0	1	2	3	rigid
13.	Euphoria	none	0	1	2	3	extreme
14.	Dependency	none	0	1	2	3	extreme
15.	Passivity	none	0	1	2	3	extreme
16.	Aggressiveness	none	0	1	2	3	extreme
17.	Hostility	none	0	1	2	3	extreme
18.	Depression	none	0	1	2	3	extreme
19.	Negativism	none	0	1	2	3	extreme
20.	Anxiety	none	0	1	2	3	extreme
IV.	INTER-FUNCTIONAL RELATIONSHIPS						
21.	Degree of insight	absent	0	1	2	3	maximal
22.	Severity of illness	mild	0	1	2	3	extreme
23.	Duration of illness	brief	0	1	2	3	extensive
24.	Level of maturity	immature	0	1	2	3	mature
25.	Interpersonal relationships	poor	0	1	2	3	excellent

FIGURE 1. RATING FORM USED FOR EVALUATION OF ALL SUBJECTS ON THE BASIS OF RORSCHACH DATA AND PSYCHIATRIC INTERVIEW

In designing the form, the authors attempted to fulfill two criteria. In the first place, the technique of assessment should not depend only upon individual categories and thus be subject to the criticism that the scoring categories had been used in a piecemeal fashion and that the total gestalt had been violated. Therefore, at every level the entire amount of data available could be utilized for making a rating on every item. Secondly, it was the intention of the authors to include personality characteristics which would be considered relevant to a psychiatrist as well as to a Rorschach expert. Accordingly every effort was made to avoid loading the scales with choice bits of Rorschachiana (such as degree of coarctation, response to inner promptings, etc.) which often resist translation into practical terms.

On the original rating form, too lengthy to permit inclusion here, the 25 variables were fully defined for the raters. A four-point scale (zero to three) was chosen in order to minimize any tendency to give average ratings for an item about which the rater felt insecure. Direction of what might be considered a "good" rating was deliberately varied, with optimity sometimes at zero, sometimes at three, and sometimes at a point in between.

Results

Total agreement among raters as influenced by levels. The first analysis made was a comparison of the amount of agreement found when three raters were all judging the same material (control group) and when level of Rorschach data was varied (experimental group). The agreement was indicated by a score derived from a system of weighting which ranged from 6 when the agreement was perfect to 1 when there was maximum disagreement. Such a weighted score was used in preference to the raw scores, since mean raw scores computed from the three individual ratings could be misleading as to the actual amount of agreement present (*e.g.*, ratings of 2, 2, 1 on a given item would receive the same mean score as ratings of 3, 2, 0, whereas the latter actually show less total conformity). No correction for chance agreement was made as this should operate equally in both groups. Table 2 presents the results of an analysis of all the items on the rating scale for all the subjects of the experimental and control groups. The average agreement score in the control group in which all raters possessed the same level of data was 4.64. In the experimental group, where level of data was systematically varied, the agreement score was 4.56. The difference between these scores is statistically nonsignificant, and any greater variability in the ratings made when different levels of data were available can be attributed to chance fluctuation.

Table 2. Comparison of Experimental and Control Groups on Total Agreement Scores.

Group	N	Mean Agreement Score	Sigma	Difference	*t*
Control	10	4.64	.46		
Experimental	24	4.56	.39	.08	.67

An analysis was made of the difference in agreement scores between the two groups for each individual item. The *t*-ratios for this analysis ranged from .03 to 2.04, with only the last figure attaining significance at the .05 level. When so many analyses are made (25), at least one difference of significance at this level would have been predicted on the basis of chance alone. This one difference occurred on Item 7, Activity Level, and was in favor of the experimental group. That is, there was greater agreement in the experimental procedure where levels were varied than there was when all raters possessed the same amount of information. There is no need to attempt to explain this logically, however, as this one difference would represent no more than chance expectation. This absence of significant differences in the individual item analyses was somewhat surprising, as on an *a priori* basis the authors had hypothesized that for certain items the content of the responses would be vital for interpretative purposes and that on these items ratings made at Level III would be hazardous.

Agreement with criterion as influenced by levels. The unit of measurement employed on this analysis was the absolute discrepancy between the ratings made by the psychiatrist (henceforth referred to as Criterion or C-Level) and ratings made at each of the three Rorschach levels. This means that the discrepancy scores could range from zero to 3 points on each item in the rating form. If at Level II the ratings became less reliable than the corresponding ratings made at Level I, then the difference between the discrepancy scores for Level C and Level I and the discrepancy between Level C and Level II should be significantly large. Exactly the same procedure could be repeated for Level III. In Table 3 the amount of total and partial agreement on all subjects in the experimental group is presented in percentages.

TABLE 3. PERCENT OF AGREEMENT BETWEEN CRITERION LEVEL AND RATINGS MADE AT DIFFERENT RORSCHACH LEVELS ON ALL SUBJECTS

	Complete agreement	1-point discrepancy	2-point discrepancy	3-point discrepancy
Criterion and Level I	37.33	46.50	15.83	.33
Criterion and Level II	35.33	47.17	16.33	1.17
Criterion and Level III	33.33	50.33	14.50	1.83

Table 4 presents the results of the analysis in terms of discrepancy scores.

TABLE 4. COMPARISON OF DISCREPANCY BETWEEN RATINGS MADE AT DIFFERENT RORSCHACH LEVELS AND OUTSIDE CRITERION (LEVEL C) ON 25 ITEMS FOR 24 SUBJECTS.

	(C - I)	(C - II)	(C - III)
Mean	19.79	20.83	21.21
Sigma	5.20	6.03	5.28
t		1.22*	2.01**

* t-ratio of difference between (C-I) and (C-II).
** t-ratio of difference between (C-I) and C-III).

From the table it is apparent that although there is no significant difference, there is a suggestive trend toward greater discrepancy as level of data is reduced, with the difference between the Level C-Level I discrepancy and the Level C-Level III discrepancy just escaping significance at the .05 level.

The individual items were also analyzed according to this procedure. In comparing Levels I and II with Level C, there is only one significant difference ($t = 2.35$) on the 25 items. This occurred on Item 4, Respect for Reality, on which Level I ratings agreed significantly more with the ratings made by the psychiatrist. In comparing Level III with Level I in terms of criterion similarity, however, three differences significant at the .05 level emerged. These occurred on Item I, Intellectual Level ($t = 2.23$), Item 2, Estimate of Intellectual Potential ($t = 2.62$), and Item 17, Negativism ($t = 2.61$). On the first two the ratings made at Level I were significantly closer to the criterion, whereas Level III agreed more closely than did Level I with the psychiatrist's ratings on Negativism.

DISCUSSION

From the above results it is apparent that the interpretative rating made on the basis of the Rorschach technique is not greatly affected by varying the level of data given the interpreter. However, it is also apparent that there is a slight tendency for ratings to coincide less closely with a criterion (psychiatric rating) when neither observation of the patient during test administration nor knowledge of the content of his responses is possible. An analysis of the individual items on which this is most likely to occur gives relatively few clues as to just what might account for this. For example, ratings on Functioning Intellectual Level and on Intellectual Potential were the two items where Level III deviated more widely from the criterion than did Level I. Yet this is an area where psychiatrists generally rely upon

psychological measures to validate their clinical judgment rather than vice versa. With respect to Negativism the ratings based on the scoring summary only (Level III) agreed more closely with the criterion than did those based upon observation of the subject in addition to the full protocol. This was particularly interesting in view of a marginal comment made by one of the raters that when rating Negativism from Level III he felt that he was doing little more than guessing.

Although one may attempt to explain the pattern of results on the basis of the items used in the rating form, this does not seem to be adequate. That is, many of the items are difficult to rate on Rorschach data alone. However, in the construction of the form, the authors kept constantly in mind the fact that ratings would be made on the basis of both psychiatric interviews and Rorschach data. Furthermore, if this were really an important factor, more differences should have appeared on the individual item analyses on those items where the Rorschach is traditionally considered more appropriate. In short, one must conclude that there is at best a slight reduction in sensitivity in using the Rorschach technique as one decreases the amount of data available for interpretation. This may either indicate that the scoring categories are reasonably valid and do succeed in quantifying the more relevant features of the test data, or else that rating on the basis of even the maximum amount of data is itself too unreliable to permit such a comparison. The use of a dual method of analysis which measures both amount of agreement and similarity to a criterion suggests that the former conclusion is more warranted.

Summary

An attempt to investigate efficiency of interpretation of Rorschach data based on the subject's behavior in its complex (Level I), the response protocol (Level II), or the scoring summary only (Level III) revealed little loss in sensitivity of the instrument as level of data was reduced. This was true whether the method of analysis was comparison of the total agreement present when all raters had the same level of information and when they had different levels, or whether similarity to an outside criterion was measured. However, there is a slight and suggestive tendency for ratings based on the maximum amount of information to agree more closely with psychiatric ratings than those based on scoring summaries alone. The results indicate that the quantitative categories into which Rorschach responses may be translated represent meaningful behavioral attributes and that, with respect to items considered relevant for this study, little information is sacrificed by relying solely on these quantitative data.

References

1. Hertz, M. R. Validity of the Rorschach Method. *Am. J. Orthopsychiat.*, 1941, 11, 512-519.
2. Ulett, G. A. Rorschach's Test as used in the Psychiatric Interview. *Northwest Med.*, 1949, 48, 544-548.

RORSCHACH TEST-RETEST AS A GAUGE OF PROGRESS IN PSYCHOTHERAPY

JOHANNA KROUT, PH.D.
and
MAURICE H. KROUT, PH.D.
Chicago Psychological Institute

and

THEODORE J. DULIN, M.D.
Chicago Psychological Institute and Psychiatric Institute, Municipal Courts of Chicago

Problem

The chief objectives of this study have been (a) to devise an experimental design and develop a system of dynamic concepts by which we could differentiate Rorschach results obtained by analytic from those obtained by non-analytic therapies; (b) to discover whether there are differences in degree and/or depth of change in analytic and non-analytic therapies; and finally (c) to determine, by way of control, whether changes in subjects needing therapy could occur without therapy and, if so, in what respects and to what extent.

Method

Psychotherapy, for our purposes, was defined as a professional relationship involving regular sessions between an individual motivated to seek help with his emotional problems and a clinical psychologist or psychiatrist following a recognized system of therapy. The psychoanalytic type of therapy was defined as the technique attempting to deal with fantasy material through free association. Non-analytic psychotherapy (of the supportive type) was defined as a technique which does not attempt to probe into the deep fantasies of the patient, and relies on the patient consciously to supply material for interpretation.

Forty two subjects, aged nineteen to thirty five, were employed in this study. The mean educational achievement of the analytic group was fourteen years. The mean educational achievement of the non-analytic group was thirteen years. The modal achievement of both groups was fourteen years. Nineteen of the subjects (14 male, 5 female) were in therapy with one of the co-authors (TJD) as psychoanalytic patients. Fourteen subjects (9 male, 5 female) were treated by two of the co-authors (MHK and TJD) by non-analytic techniques. Nine subjects (5 male, 4 female), though diagnosed as in need of psychotherapy, failed to obtain it, but returned for Rorschach retest at a later date. These untreated cases, serving as controls had certain characteristics in common: (1) They manifested clinical anxiety at the time of testing; (2) their tests indicated need for therapy; and (3) a subsequent crisis, or series of crises, outside the framework of therapy as here defined, occurred in each case. The general composition of the three test groups (analytic, non-analytic, and untreated) was well balanced from the point of view of therapeutic need. The neuroses made up 70 to 85% of each of the groups, while pre- and post- psychotics constituted the remainder. A state of balance existed likewise in regard to the average periods elapsing between test and retest. The analytic cases averaged 12 months in therapy, and totaled an average of 150 hours during that period. The non-analytic cases spent an average of 13 months in therapy, totaling an average of 40 hours. In the non-treatment cases an average period of 24 months elapsed between Rorschach test and retest. It was established that such events as marriage, divorce, separation, failure at school, leaving home, sex experiences, etc., occurred in each case, motivating the subject to re-apply for service.

Test Procedure

It has been the custom among Rorschachians to speak in terms of a molecular as against an atomistic type of protocol interpretation. The former indeed has served as the basis for most clinical reports. Yet, the current literary output of

Rorschach investigators, and especially some recent statistical studies popular among non-clinicians, show a tendency to revert to the segmental or atomistic type of analysis. The present study is an attempt to report on Rorschach findings in terms which experienced Rorschach clinicians have actually used for some time, but which in the literature have not received proper emphasis. To be specific, clinical experience with the Rorschach has demonstrated that the qualities of personality reflected in the test do not inhere in any one scoring category. For one thing, certain scoring categories may not appear in a particular protocol. Furthermore, the significance of a particular category as related to other categories necessarily varies with each case. For this reason, and to make global evaluation of Rorschach material not only more defensible but actual, we developed a system of dynamic concepts which are of especial interest clinically, and applied this system to each of the eighty four protocols with which we were concerned. The system of concepts, and the scoring items subsumed, appear in Table 1.

TABLE 1. A SCHEMA FOR CLASSIFYING RORSCHACH MATERIALS*

1. ANXIETY
   - Diffuse shading Rs
   - RT
   - Card-turning
   - Incidental remarks (especially requests for evaluation, etc.)
   - Evasive Rs
   - Dd%
   - (H + A) : (Hd + Ad)
   - Shading shock
   - Etc.

2. LIBIDO
   - Color Rs
   - Symbolic content of Rs (especially cards II, III, & VI)
   - Nature of movement Rs
   - Etc.

3. HOSTILITY
   - m
   - C
   - S
   - Symbolic content of Rs
   - Nature of movement Rs
   - C'_w, c
   - Etc.

4. GUILT
   - Ch distribution
   - Symbolic content of Rs
   - Ch:C ratio
   - Etc.

5. INTERPERSONAL RELATIONS
   - H%
   - Number and kind of M
   - Number P
   - Etc.

6. INTEGRATION
   - At
   - F%
   - W%
   - Re%
   - Shading shock
   - Color shock
   - Etc.

7. APPROACH
   - M:C
   - (Fm + m) : Ch
   - Re%
   - %W, %D, %d, %Dd/S
   - O%
   - Number P
   - Etc.

8. REALITY TESTING
   - F+%
   - FK
   - k
   - M:C
   - F%
   - FM%
   - Etc.

9. PATHOGNOMONIC SIGNS
   - Midline references
   - Contaminations
   - Perseverations
   - Confabulations
   - Positionals
   - Disembodiments
   - Bizarre associations
   - Stereotypy
   - Etc.

10. PRODUCTIVITY
    - Number R
    - Re%
    - W:M
    - O %

*The particulars subsumed under the ten concepts in this schema are in themselves not new, and are meant to be suggestive rather than exhaustive. The system of concepts, and the range of items grouped under them, constitute what is new in this presentation. Each competent Rorschach worker, drawing on the literature and his own clinical experience, could add or detract from the list of particulars.

Essentially, our procedure was as follows: The subjects employed in this study presented themselves for diagnostic interview in the office of a psychiatrist or a clinical psychologist, and were at once referred for Rorschach testing. The subjects were re-tested (as previously indicated, 12 to 13 months later) by the same person in each case, and the two protocols obtained were evaluated on the basis of the concepts in the system employed.

In terms of our experimental design, each pair of Rorschach protocols was evaluated by one of the co-authors who had had no previous familiarity with the cases, and was not familiar with the type of therapy employed. This individual judged the protocol as "changed" or "unchanged" in each concept area. His judgments were then matched against the judgments of another co-author who similarly evaluated a sample of the protocols. A high order of consistency was achieved in spite of the fact that, in global analysis as we have used it, precise weighting of each scoring entity is not possible.

Results

The test-retest material for each of the forty-two subjects was grouped according to presence and direction of change. The chi square test, with correction for small samples, was applied as a measure of significance (Table 2.)

Table 2. X^2 and P Values

RORSCHACH SCORING CLASSIFICATION	Analytic Cases		Non-Analytic Cases		Untreated Cases	
	X^2	P	X^2	P	X^2	P
ANXIETY	.4	.5	.06	.8	.2	.70
LIBIDO	6.1	.015	9.0	.01	3.0	.07
HOSTILITY	6.1	.015	1.2	.25	3.0	.07
GUILT	7.8	.01	1.2	.25	.2	.70
INTERPERSONAL RELATIONS	6.1	.015	7.0	.01	.9	.40
INTEGRATION	6.1	.015	.3	.5	.9	.40
APPROACH	7.8	.01	2.0	.15	5.5	.02
REALITY TESTING	.4	.8	.3	.5	.2	.60
PRODUCTIVITY	.4	.8	.3	.5	5.5	.02
PATHOGNOMONIC SIGNS	2.0	.15	7.0	.01	.9	.40

Degree of freedom in all cases equals 1.
Correction for small samples equals 0.5.

The analytic cases showed significant change, at the one percent level, in the concept-areas of Libido, Hostility, Guilt, Interpersonal Relations, Integration, and Approach. The non-analytic group showed change, at the one per cent level of confidence, in the concept-areas of Libido and Interpersonal Relations. In the untreated (control) groups, we found significant changes at the one per cent level of confidence in Approach and Productivity.

The statistical significance of these findings underscores the importance and reliability of the several findings in each of the three groups. The fact that statistical significance is lacking in four area-concepts (Anxiety, Reality Testing, Pathognom-

onic Signs, and Productivity) in the analytic group; in eight area-concepts (Anxiety, Hostility, Guilt, Integration, Approach, Reality Testing, Pathognomonic Signs, and Productivity) of the non-analytic group; and in eight area concepts (Anxiety, Libido, Hostility, Guilt, Interpersonal Relations, Integration, and Reality Testing) of the untreated group, does not mean, of course, that some of the individuals in each of the three groups did not change considerably in any of the remaining areas. It simply means that the changes that did occur in other areas could not reasonably be expected to occur beyond the range of chance; and hence could not be said to offer grounds for accurate prediction.

From the results obtained, we can see that either of the two types of therapy can apparently effect statistically significant changes in such areas as Libidinal Organization and Interpersonal Relations.This may mean that these two categories are interdependent; or that change in the one implies change in the other. It may even be that they can be plotted on a single continuum, and thus represent various degrees of one basic quality.

Among the concepts that have proved significant for the analytically treated group, we find Hostility and Guilt. It is of interest to note that another concept, Integration, is also found significant for the same group. The hypothesis is advanced that, possibly, change in Integration is a function of change in Hostility and Guilt; and hence these three categories tend, for that reason, to occur together with equal frequency.

The category of Approach does not lend itself to a similar hypothesis, since it does not seem to depend on therapy at all. It will be recalled that it was found with significant frequency in both the analytically treated group and in the totally untreated group. The conclusion this suggests is that change in approach is independent of change in other areas included in the system. For this reason it is highly probable that the area of Approach, as well as the area of Productivity, are the only areas that changed without therapy in those subjects with serious emotional problems. This does not mean of course that personalities cannot be modified in other ways, but that they do seem to be modified more consistently, and perhaps more easily, in these areas than in any of the others we employed.

It is of some interest to note that Libidinal Organization and Interpersonal Relations were the only areas showing significant change in non-analytic subjects. This may mean that the findings of this study are essentially a function of the intensity of therapy, rather than of differences in technique employed. Since the non-analytic cases received (in terms of hours) less intensive treatment, it may be that shifts in these two areas occur—in a way discernible in Rorschach data—*before* other areas begin to shift. This need not mean that there is not a continuous change in libidinal organization beyond a certain point in therapy. It does mean that Rorschach begins to show it at a certain point. What is most interesting, of course, is that Libidinal Organization tends to change at all without so-called "deep" interpretation.

Comment should be made on the categories of Anxiety, Reality Testing, and Pathognomonic Signs which did not show statistically significant change in *any* of the three experimental groups. Reality Testing and Pathognomonic Signs are taken to refer to degree of perceptual distortion. In dynamic personality theory, it is held that perceptual distortion is traceable to an anxiety disturbing the Ego. Furthermore, anxiety in the personality (conceived in a total, quantitative sense) is an index of unresolved conflicts. It appears natural then for the category of Anxiety, as we dealt with it, to be *last* to show change in the ongoing therapeutic process, since the uncovering of deeper material always brings with it fresh onslaughts of anxiety upon the Ego. At the point where personality equilibrium is finally achieved, the total amount of manifest anxiety may be expected to be radically decreased, and the reality picture to become clearer to the subject at the same time. It is at this point, probably, that anxiety nuances become discernible on the Rorschach, and it is at this point, also, that reality testing becomes sufficiently intense to reduce pathological signs in proportion.

The system of analysis employed was limited entirely, of course, to Rorschach signs. It is another problem to understand the process, and interpret the significance, of registering specific variables in Rorschach responses. Other changes may be seen clinically and through other tests. Furthermore, representation of change reflected in the Rorschach may not necessarily be present clinically. Clarification of these issues will eventually enrich our understanding not only of clinical processes but of personality generally. The findings of this paper do show, however, that the Rorschach technique reliably distinguishes personality changes in three groups of subjects, through the use of molecular, rather than atomistic, interpretations in areas meaningful to the therapist.

Conclusions

The experimental design involved the use of Rorschach test-retest with nineteen subjects undergoing psychoanalytic therapy, fourteen subjects undergoing non-analytic (supportive) therapy, and nine untreated subjects who were in need of therapy. Regarding the analytic subjects as our experimental group, we used the non-analytic and untreated groups as controls. The pathological status of the three groups as reflected in diagnosis was approximately equal in percentage terms. The average duration of therapy (not therapeutic hours) was almost exactly the same in both therapeutic groups; though in our untreated group time elapsing between Rorschachs was twice as long as it was in either of the therapeutic groups.

Our findings indicate that personality changes were more extensive, and probably more profound, in the analytic than in our supportive-therapy groups, though these changes could be explained in terms of intensity of treatment and the time variable (session-hours) involved. The therapeutic-control, or non-analytic, group showed less thorough modification of subjects' personality dynamics, and did not show any changes which were not also significantly present in the analytically treated group. For the untreated controls, changes were obtained in two areas, only one of which was also present in a therapeutic group. In a negative way, this finding tends to validate the significance of the changes in our therapeutic groups. No statistically significant changes were found in three concept-areas included in our system, and the explanation offered is that these are coordinate concepts whose modification must be expected, if at all, on even deeper levels of therapy than either of our therapeutic groups were afforded in the time involved.

THE DISCRIMINABILITY OF THE "DRIVE FACTORS" REPRESENTED IN THE SZONDI PICTURES

HELEN M. RICHARDSON

New Jersey College for Women, Rutgers University

Problem

One of the basic Szondi assumptions is that the six photographs of patients in a given diagnostic category express a common "need-system" or "drive factor" which is present in exaggerated form in that type of mental disturbance. The significance of a subject's choices is derived from his having discriminated and responded to this significant common feature. The individual picture in the category does not matter; it is the number of choices from the category that counts. From this, it appears to follow that the six pictures belonging to a given category should be perceptibly alike with respect to this common character. They should be discriminable on this basis from the pictures belonging to other diagnostic categories. Furthermore, this common characteristic should be the most response-compelling thing about each picture. An attempt to match together the pictures which look as though they expressed the same drive factor should therefore result in matchings that correspond to the diagnostic categories. The problem of this study is to find out how far this is actually the case.

Subjects and Method

The subjects were four men and six women of high average or superior intelligence. Four were between 20 and 25 years of age; two in the fifties; the other four between 25 and 50. Three were college undergraduates. Three served as subjects as a favor to the experimenter; the other seven were paid at the standard rate for student non-technical assistants. One of the ten subjects performed the experimental series twice. Only forty per cent of her matchings duplicated each other, and both of her series have been included in the count of matchings. Thus it is as if there were eleven subjects.

The method was as follows: The eight cards of Set I were laid out on a table in front of O (the subject), about twelve inches from the table edge. The cards were laid out in the arrangement prescribed by Deri [1, p. 8] for giving the Szondi test, so that they would be seen in the standard configuration. While laying out Set I, E (the experimenter) said:

> "Here is a set of pictures of eight people. Each of these persons may look to you as if he had some dominant emotion or attitude toward life or some characteristic way of feeling about himself and other people. Or he might look as if he were likely to act in a certain way."

Set II was then laid out after the standard pattern, but on O's right and near the edge of the table. E continued:

> "Now here is another set of eight pictures. Try to match each of these pictures with the one in the first set that shows the same dominant emotion or attitude or way of feeling. Arrange the pictures in the second set so that each one is in the position corresponding to the corresponding picture in the first set."

When O indicated that he had finished, E removed Set II, leaving Set I on the table. Set III was then presented. O was told that there would be five sets to match with Set I, and that when these were finished Set II would be set up as a standard to be matched in turn by each of the other sets. Following this pattern, each of the six sets served its turn as standard as well as coming up to be matched against each of the other sets. Each set was therefore twice exposed to matching with each other set.

No time limit was set, but E recorded the time as each of the sets was laid out as a standard, and when the last matching was completed. While O was working on a given set, E recorded the results of the preceding matching. Also E noted down, as far as possible, comments or remarks made by O. No explanation of the pictures was given by E other than that the same eight dominant emotions or attitudes were supposed to be represented in each of the six sets. One O, a professor of psychology, knew the nature of the test but was not familiar with the pictures.

After O had completed the final matching, with Set VI as standard, E continued:

> "Now I want to see whether there is any relation between the kind of picture that a person is best at matching and the way he feels about that kind of people—whether he likes them or dislikes them or doesn't care one way or the other."

With this introduction, E went on with the standard procedure for administering the Szondi test [1, pp. 9-12].

Results and Discussion

The results of the matchings were tabulated on a grid with the eight diagnostic categories arranged from left to right and from bottom to top after the plan of Table 1. Within each category the pictures from I to VI were assigned spaces in similar fashion. It was thus possible to see what individual pictures were matched with each other, as well as to note the matching by categories.

TABLE 1. DISTRIBUTION OF CATEGORY-CATEGORY MATCHINGS AMONG THE SZONDI PICTURES.

Figures in parentheses are based on totals from the two cells in which each combination is represented. In cells which represent matching of category with itself, a comparable result is attained by counting the single frequency twice. Percentages are based on the total of 660 matches for each category. Chance expectancy is 12.5%.

	Homosexual	Sadistic	Epileptic	Hysteric	Catatonic	Paranoid	Depressive	Manic	Total
Manic	28 (58) (9%)	44 (85) (13%)	25 (58) (9%)	14 (29) (4%)	64 (121) (19%)	18 (34) (5%)	25 (51) (8%)	112 (224) (34%)	330
Depressive	27 (69) (10%)	22 (37) (6%)	50 (103) (16%)	64 (127) (19%)	48 (92) (14%)	39 (73) (11%)	54 (108) (16%)	26	330
Paranoid	34 (65) (10%)	35 (77) (12%)	42 (82) (12%)	35 (68) (10%)	56 (105) (16%)	78 (156) (24%)	34	16	330
Catatonic	7 (15) (2%)	29 (50) (8%)	32 (60) (9%)	61 (115) (17%)	51 (102) (15%)	49	44	57	330
Hysteric	15 (42) (6%)	44 (73) (11%)	54 (102) (15%)	52 (104) (16%)	54	33	63	15	330
Epileptic	49 (91) (14%)	39 (84) (13%)	40 (80) (12%)	48	28	40	53	33	330
Sadistic	63 (106) (16%)	74 (148) (22%)	45	29	21	42	15	41	330
Homosexual	107 (214) (32%)	43	42	27	8	31	42	30	330
Total	330 (660)	330 (660)	330 (660)	330 (660)	330 (660)	330 (660)	330 (660)	330 (660)	2640 (5280)

Table 1 shows the frequency with which each category was matched with itself or with each other category by all subjects combined. The figures without parentheses are based on the actual tabulations which covered the entire grid. Reading the figures *across the row* labeled *Manic*, we find the number of times that a *manic* picture as a standard *received* matchings from the pictures in the eight diagnostic categories, including its own. Reading *down the column* labeled *Manic*, we find the number of times that *manic* pictures were *assigned as matches* to pictures in the eight categories. The figures in parentheses combine the column with the row to show the *total* number of matchings of *manic* pictures among the eight categories. In the totals, the frequencies for matching of a category *with itself* have been doubled, since they must be counted both as rows and as columns. Counted thus, there are 660 matches for each category. Chance expectancy for a given picture to be matched with a given standard is 12.5 per cent.

Matchings of *manic* pictures with *manic* pictures and of *homosexual* with *homosexual* occur with almost three times chance frequency. A frequency almost double the chance expectancy appears for matching of *paranoid* with *paranoid* and of *sadistic* with *sadistic*. In these four categories, the self-matchings are significant at the .01 level. Even at this, however, they seem to fall short of the precision implied in Szondi test interpretation. In the *epileptic*, *hysteric*, *catatonic*, and *depressive* categories, self-matchings do not deviate significantly from chance expectancy.

Table 1 is of some interest also for its revelation that there is no category which is not at some time matched with each of the other categories. True enough, this may be due to the fact that the matching was forced into a set and limited pattern. Every wrong match forced at least one other wrong match. The matchings which have a frequency of less than half the chance expectancy may be worth noting. Matchings of *catatonic* with *homosexual*, of *manic* with *hysteric* or *paranoid*, and of *depressive* with *sadistic* fall into this distinctly low frequency class.

The total number of same-category or "correct" matchings, for all categories combined, amounts to 22 per cent of the total of 5280 matchings. Scores of individual subjects ranged from 15 per cent correct to 29 per cent, with the median at 21.5. The lowest individual scores do not differ significantly from chance. The time taken for the matching ranged from 62 to 137 minutes. Rank-difference correlation between score and length of time is −.44.

More meaningful than the category matchings of Table 1 are matchings between individual pictures. In this analysis we draw on the detailed chart of tabulations, which it is not feasible to set up in type. Our experimental procedure (with 11 subjects) permitted a maximum of 22 matchings of each individual picture with each other picture (except, of course, that pictures *within the same set* were not subject to matching with each other). In Table 2 are listed all the pictures which were paired in at least 50 per cent of the opportunities that were offered. The difference between this percentage and the chance expectancy of 12.5 per cent is significant at the .01 level. There are 32 such pairings. (In Tables 2 and 3 and in the discussion following, the categories are designated by their initial letters, with "k" for catatonic, as they are on the back of the Szondi cards.) Ten of the pairings listed in Table 2 are between same-category pictures: 4 involving matching of an *m* picture with an *m* picture, 3 of *h* with *h*, and one each of *p* with *p*, of *k* with *k*, and of *s* with *s*. The other 22 are pairs from different categories.

In Table 3 are listed all the *same-category* pairs that were matched by our subjects with *no more* than chance expectancy; that is, in not more than 3 out of the 22 opportunities.

From Tables 2 and 3 we see that the relatively large number of correct matchings in the *manic* category which showed up in Table 1 was due chiefly to the resemblance among the *m* pictures in Sets II, III, and V. The *m* and *k* pictures in Set VI seem to bear no kinship to the other pictures in their respective diagnostic categories. It appears that Szondi subjects who choose the *k* picture in Set VI may be identifying themselves with *m* drives rather than with *k* drives. Similarly, choice

TABLE 2. MATCHINGS BETWEEN INDIVIDUAL SZONDI PICTURES WHICH OCCURRED WITH A FREQUENCY AT LEAST FOUR TIMES AS GREAT AS THE CHANCE EXPECTANCY

Frequency of Matching		Matched Pairs	
N	%	Same Category (10 pairs)	Different Categories (22 pairs)
22	100		
21	95		*k*VI *m*III
20	91		*k*VI *m*II
19	86	*m*II *m*V	
18	82	*m*II *m*III	
17	77	*m*III *m*V	*k*VI *m*V
16	73		*k*VI *s*I
15	68	*h*III *h*IV *p*II *p*V	*k*II *d*III *k*II *hy*I
14	64	*m*I *m*II *s*I *s*II	*e*VI *m*I *e*VI *s*II
13	59	*k*I *k*III	*k*VI *m*IV *h*III *s*II *h*III *s*VI *hy*I *d*III
12	55		*m*III *s*I *e*II *d*I *s*V *p*IV
11	50	*h*II *h*V *h*V *h*VI	*h*I *p*IV *h*V *s*III *e*IV *p*II *e*I *d*VI *d*VI *p*III *p*V *k*III *k*V *d*III

TABLE 3. SAME-CATEGORY PICTURES WHICH WERE NOT MATCHED WITH EACH OTHER

Category	Pictures which were not matched together
h	
s	
e	*e*I with *e*II, *e*IV, or *e*VI; *e*IV with *e*II, *e*III, *e*V, or *e*VI
hy	*hy*I with *hy*III
k	*k*VI with *k*I, *k*II, *k*III, *k*IV, or *k*V; *k*II with *k*III
p	*p*IV with *p*I or *p*VI
d	
m	*m*VI with *m*I, *m*II, *m*III, *m*IV, or *m*V

of *e*VI would seem to be more indicative of identification with the drives expressed in *m*I or *s*II than with *e*I or *e*IV. *k*II has more in common with *d*III and *hy*I than with *k*III or VI, and *hy*I is more like *d*III than like *hy*III. True enough, there are a number of cases where same-category pictures are judged to show "the same dominant emotion, attitude, or way of feeling." But *more common* are instances of high-frequency judgments of similarity of drive between pictures from *different* diagnostic groups.

TABLE 4. PREFERENCE SCORES OF THE SIX INDIVIDUAL PICTURES IN THE EIGHT DIAGNOSTIC CATEGORIES

Preference Score	Manic	Homosexual	Paranoid	Sadistic	Epileptic	Catatonic	Depressive	Hysteric
10	mII							
9								
8	mI	hV	pIII					
7								
6	mV							
5				sIV				
4	mIII	hVI	pIV		eI, eVI	kIV		
3	mVI			sI	eIII	kVI	dII	
2		hII						
1			pV	sV				hyI
0					eIV	kII	dIII	
–1		hI, hIII	pII	sII, sIII				hyV
–2	mIV			sVI				hyIV
–3			pI				dV, dVI	
–4		hIV	pVI			kV	dI	
–5						kIII	dIV	hyVI
–6					eII			
–7								hyII
–8					eV	kI		hyIII
Total Score for Category	29	8	5	5	–3	–10	–12	–22

The attempt to discover whether there was any relation between an individual's success in matching within a given category and that subject's likes and dislikes for the pictures yielded no significant results. However, these expressions of likes and dislikes provide some data of positive value. The choices reveal wide differences in popularity among the individual pictures in a given category as well as among the eight categories. A crude "preference score" for each picture was derived by taking the algebraic sum of all the positive and negative choices received by that picture from all subjects. These scores range from + 10 for *m*II to −8 for *e*V, *k*I and *hy*III. Our findings regarding the order of popularity agree closely, though not perfectly, with some results reported by Harrower[2]. After conversion of Harrower's data into preference scores by the method described above, the correlation (rho) between the two sets of scores was found to be .88.

Table 4 shows the preference scores for the six individual pictures in the eight diagnostic categories, arranged so as to show the range in popularity of the pictures within a category and the difference in total popularity among the categories. If we arrange the actual pictures in the order indicated by Table 4, and refer again to Tables 2 and 3, we can see that degree of smiling, of appearing friendly, sociable, or competent, *vs.* withdrawn, suspicious, or anxious, may well have been the kind of "emotion, attitude, or feeling" on the basis of which our subjects matched the pictures and also the basis for their preferences. In some cases, resemblance in these characteristics occurs between pictures within the same diagnostic category; in some cases it occurs between pictures in different categories.

Summary

Tests were made to find out to what extent Szondi pictures having the same diagnostic classification would be judged to express the same "drive factor." Correspondence appears essential to the assumptions underlying the Szondi test. Ten subjects twice matched the eight pictures of each set with those of each other set, arranging in corresponding positions pictures that represented "the same dominant emotion, attitude, or way of feeling." After the last series of matchings, each subject was asked to indicate the pictures liked best and least from each set.

Chance expectancy of matching any diagnostic category with any category (including its own) would be 12.5 per cent. The percentages of correct (same-category) matchings were as follows: *h*:32; *s*:22; *e*:12; *hy*:16; *k*:15; *p*:24; *d*:16; *m*:34. Four of these are significant at the .01 level; the other four are below the .05 level of significance.

The frequency with which each individual picture was matched with each other picture was also noted. There were 32 pairings between individual pictures which occurred in 50 per cent or more of the opportunities provided by the experimental procedure (significant at the .01 level). Of these 32 pairings, 10 were between same-category pictures; 22 were between pictures from different categories. The *k* picture from Set VI was consistently matched with *m* rather than with *k* pictures.

Choices of pictures liked best and liked least disclosed no significant relation between factor loading and success in matching of category with category. It appears, however, that degree of looking friendly, sociable, competent, vs. withdrawn, suspicious, or anxious may have been a basis for both the choices and the matchings.

References

1. Deri, Susan. *Introduction to the Szondi Test.* New York: Grune & Stratton, 1949.
2. Harrower, Molly R. Experimental studies with the Szondi Test. Psychological Service Center, Univ. Okla., (Mimeo.)

A RORSCHACH STUDY OF THE PERSONALITY STRUCTURE OF OBESE WOMEN*

BENJAMIN KOTKOV AND BENJAMIN MURAWSKI

U. S. Public Health Service *Boston VA Mental Hygiene Unit*

Introduction

Currently, there is a wave of interest in the problems of the obese individual. The literature discloses that a study of the obese personality is important because of the high incidence of the disorder[4], its serious physical and physiological correlates[4, 8] and a large number of studies[1, 2, 3, 5, 11] which have reported or suspected a relationship between obesity and more or less specific personality dimensions.

The purpose of the present study is to approach the adult female obese personality through the use of the Rorschach test.[1] The authors' plan is (1) to determine if there are reliable and differentiating Rorschach quantitative indices which characterize the personality structure of obese women as contrasted with women within the ideal weight range, (2) to determine the degree in which these differentiating indices are present, and (3) to attempt to reconstruct a pattern of psychic-economic distribution of personality energy typical of the personality structure of obese women. Later papers will treat nuclear conflicts and crucial relationships in the personality of obese women as reflected in other projective techniques.

Methods

The subjects were volunteers in the Boston Experiment for Group Psychotherapy with the obese[6] sponsored by the Chronic Disease Division of the U. S. Public Health Service in cooperation with the Massachusetts State Health Department. The obese population was composed of volunteers who responded to newspaper publicity and a radio program sponsored by the Massachusetts Society of Mental Hygiene. All applicants had to present a physician's certificate for medical clearance. Sevringhaus'[10] definition of obesity was used: "... obesity is an excess of fat over the normal expected for height, age, sex, between the ages of 25 and 30 ... an allowance of 10% variation above and below the average is to be made." The "ideal weights" or controls were volunteers from a School of Public Health, Public Health Personnel, a Visiting Nurses Association, Dispensary Personnel, Adult Education groups, a YWCA, and a Community House. Ideal weight was determined from a nomograph distributed by the Metropolitan Life Insurance Company. In view of the rather large number of obese women who were tested it was felt we could best control sampling errors by dividing our population into a "pilot group" and an "experimental group" wherein the results from the pilot group could be checked with a second and larger experimental group and those results identified which occurred by chance. This gave us two distinct populations of 51 female subjects for a pilot study and 80 female subjects for the experimental group.

The pilot group of 51 subjects were randomly selected from 131 volunteers, and from a group of 30 obese and 21 ideal weights we established preliminary hypotheses. In evaluating the statistically significant variables, however, the use of random selection from the total population tested, rather than matching for age, IQ, education, marital status and career *vs.* housewives left us in the position of not being able to say that obesity, our main concern, was the variable which produced the differences observed in our statistics.

*A paper presented at the American Psychiatric Association Regional Research Conference in Boston, February 28, 1952.

[1]A group technique (7) was used as a more convenient, expedient device. Conditions of administration were equal, hence comparable for our purposes.

The difference between the means of age, education, marital status, and career *vs.* housewives of the two groups of the pilot study were far beyond chance so that along these dimensions our two groups varied significantly. An inspection of Table 1 shows that the difference in age, education, marital status, and housewives *vs.* career women (nearness to the refrigerator) could account for the results.[2] Since our design involved splitting the parent population into pilot and experimental groups the differences created by random selection were not fatal but required correction.

TABLE 1. CHARACTERISTICS OF PILOT GROUP.

	Obese N = 30				Ideal N = 21		
Factors	Mean	Sigma	Range	Prob. Value	Mean	Sigma	Range
Age	42.82	9.87	24–63	.02	34.24	13.73	20–55
Ed.	11.63	2.12	8–18	.01	14.71	3.98	12–20
I. Q.	115.64	24.69	96–137	.40	122.00	28.84	105–135
M. S.	16 married (53%) 14 single (47%)			<.001	2 married (10%) 19 single (90%)		
Occupation	15 career (50%) 15 housewives (50%)			<.001	21 career (100%) 0 housewives (0%)		

TABLE 2. CHARACTERISTICS OF EXPERIMENTAL GROUP.

Factor	Obese N = 46				Ideal N = 34		
	Mean	Sigma	Range	Prob. Value	Mean	Sigma	Range
Age	34.41	7.89	20–50	.50	35.44	9.30	19–50
Ed.	13.15	1.90	11–18	.50	13.38	2.09	9–17
I. Q.	118.88	32.43	92–140	.50	119.65	13.52	95–134
M. S.	27 married (58%) 19 single (42%)			.70–.80	22 married (65%) 12 single (35%)		
Occupation	22 career (48%) 24 housewives (52%)			.70–.80	14 career (41%) 20 housewives (59%)		

The probability values of the means in Table 2 demonstrate clearly that the loose ends of random selection were corrected by a close matching of our subjects in the experimental group. This increases greatly the probability that any difference we can demonstrate on the Rorschach can be attributed to weight for the other controllable dimensions have been equated.

The pilot sample was subjected to extensive statistical scrutiny using chi-square and *t*-tests. The latter was employed for those variables involving percentages where the mean percent represented the most efficacious manner of dealing with the Rorschach variable. Seventeen general Rorschach formal and content categories

[2]The authors wish to acknowledge the statistical advice of Professor George Miller of the Massachusetts Institute of Technology.

were subjected to chi-square analysis, but by using a refined breakdown a total of 46 variables were tested. The range was from "number of responses per card," to "consistency of M:Sum C, FM:Fc, R to last three cards," to "type of W response." Eleven *t*-tests were run: W%, D%, F%, expanded F%, F plus %, A%, % to last three cards, R total, IQ, Age and Education. The probability value of .20 level of confidence was utilized as a cut off point in the pilot group for the construction of hypotheses. As the direction of the difference between the obese and the ideal weights was to be predicted from the hypotheses, a one-sided test of significance was employed. This allowed us to posit trends of .10 to .06 in our experimental group.

Results

Fifty-seven empirical hypotheses were considered. Chi-square comparisons yielded nine specific Rorschach hypotheses for obesity: M less than two, greater sum C than M, no recreation content, less H, more CF than FC, no botany, no clothing, no lamp responses, and more F– responses. The *t*-tests contributed five additional hypotheses: W% less than 36.46, D% more than 56.17, F% more than 60.83, expanded F% more than 62.83, and A% more than 41.30.

However, when these fourteen hypotheses, derived from the pilot group, were applied to the new population of the experimental group, seven hypotheses emerged

Table 3. A. Chi-Square Comparisons

Rorschach Factors	Pilot Group (Obese N = 30; Ideal N = 21)			Experimental Group (Obese N = 46; Ideal N = 34)		
	Chi-square values	Degree of freedom	Level of confidence	Chi-Square values	Degree of freedom	Level of Confidence (½P)
M	10.82	2	.01	4.87	2	.04
M:sum C	8.30	2	.02	1.47	2	.12
Recreation	11.45	1	.01	10.15	1	.001
H	7.62	1	.01	1.78	1	.09
FC:CF	6.86	2	.05	0.53	2	.40
Botany	2.72	1	.10	1.44	1	.12
Clothing	2.27	1	.10	2.02	1	.07
Lamp	1.69	1	.20	4.38	1	.02
F–	1.53	2	.20	4.25	2	.06

B. "*t*" Test Comparisons

	"t"	Level of Confidence	"t"	Level of Confidence (½P)
W%	2.68	.01	3.29	.001
D%	2.00	.05	3.60	.001
F%	1.61	.10	2.17	.015
Exp. F%	2.74	.01	2.85	.005
A%	1.93	.05	1.06	.17

at the .05 level of significance or higher, and three additional hypotheses indicative of trends at the .10 level or higher. The following seven Rorschach variables survived the double weeding out processes and can be considered to be significantly different for the two groups: no lamp content, no recreation content, F% more than 60.83, expanded F% more than 62.83, M less than two, W% less then 36.46, D% more than 56.17. A further inspection of Table 3 disclosed three additional trends (between .06 and .10 level of significance): less H, more F−, and no clothing responses.

There were several significant negative results as well. First, the mean number of responses was the same. This allowed us to proceed statistically much more freely than if we had to equate for the number of responses. Second, the color ratios and percentages did not hold up, removing a possible discriminating problem out of the area of relationships with external objects.

Table 4 serves to integrate the discrete variables into a composite picture of 50 subjects selected at random: 25 obese and 25 ideal weights. The significance of each variable was previously determined in isolation and since the Rorschach scores are never treated in isolation, but rather interpreted in the light of all other categories, the authors hesitated to deal structurally with the 7 variables without some measure of the frequency of occurrence in a sampling of both groups. Table 4 shows that F% and expanded F% (sum of F, FK, and Fc%) are perfectly correlated for all 50 cases. Since the retention of both signs would not explain more of the variance, F% is deleted for the purpose of arriving at a statistical index of obesity. An average of 3.7

TABLE 4. THE FREQUENCY OF RORSCHACH SIGNS WITH BODY WEIGHT.

A. OBESE SAMPLE (N = 25)

Case	Lamp None	Recreation None	F% more than 60.83	Expanded F% more than 62.93	M less than 2	W% less than 36.46	D% more than 56.17	# of Signs	% Overweight
1.	x	x						2	36
2.	x	x	x	x	x			5 (4)	11
3.	x	x				x		3	53
4.	x		x	x	x	x	x	6 (5)	53
5.	x					x	x	3	43
6.	x	x	x	x	x			5 (4)	91
7.	x	x				x		3	24
8.		x						1	34
9.	x	x						2	57
10.	x	x			x			3	34
11.	x		x	x		x	x	5 (4)	43
12.	x		x	x	x	x	x	6 (5)	16
13.	x	x	x	x	x		x	6 (5)	24
14.	x	x	x	x		x	x	6 (5)	28
15.	x	x				x	x	4	36
16.	x					x	x	3	14
17.	x	x	x	x		x	x	6 (5)	19
18.	x	x	x	x				4 (3)	25
19.	x	x	x	x				4 (3)	25
20.	x	x	x	x	x	x	x	7 (6)	27
21.	x	x	x	x	x	x	x	7 (6)	29
22.	x		x	x	x	x	x	6 (5)	30
23.	x	x	x	x				4 (3)	33
24.	x	x						4	49
25.	x	x				x	x	4	60
Totals	24	19	14	14	9	14	13	21/25 had 3 or more excluding F%	

B. Control Sample (N = 25)

Case	Lamp None	Recreation None	F% more than 60.83	Expanded F% more than 62.93	M less than 2	W% less than 36.46	D% more than 56.17	# of Signs	% Overweight
1.	x		x	x			x	4 (3)	−7
2.	x							1	−2
3.	x							1	−9
4.	x							1	6
5.								0	9
6.								0	−6
7.	x	x						2	−10
8.	x							1	0
9.	x							1	10
10.								0	0
11.	x		x	x		x		4 (3)	10
12.	x					x		2	6
13.	x	x						2	−6
14.								0	5
15.		x						1	10
16.	x							1	−10
17.			x	x		x		3 (2)	−7
18.	x							1	3
19.	x							1	0
20.	x					x	x	3	−6
21.								0	2
22.	x							1	2
23.	x					x	x	3	−4
24.	x							1	−7
25.								0	0
Totals	17	3	3	3	0	5	3	4/25 had 3 or more excluding F%	

signs (out of 6 possible signs) was obtained for the obese, and an average of 1.2 signs was obtained for the controls. This is an average of 2.46 for the two groups together. Using the combined mean as a cutting point, the 50 cases were divided into those having 2 or less versus those having 3 or more signs. The appropriate chi-square is very significant (chi-square =23.12, with one degree of freedom, level of confidence far beyond the .001 level). Hence it can be concluded that the occurrence of 3 or more of these 6 signs in the record of a female subject is strong Rorschach evidence that she has the obese personality structure.

In order to obtain the most inclusive personality description, an inspection of Table 4 was made for consistent patterns. Of the obese, cases 2, 4, 6, 11, 12, 13, 14, 15, 17, 20, 21, 22, and 25 (13 cases) have 4, 5, or 6 of the 6 signs (no controls have more than 3). Cases 4, 12, 13, 14, 17, 20, 21, and 22 (8 cases) have 5 or 6 of the 6 signs and 6 of these 8 extreme cases show lamp none, expanded F% greater than 62.93, M less than 2, and D% greater than 56.17.

Discussion

An inspection of the Rorschach variables suggested that the obese women registered the same number of scorable stimuli entering the perceptual system of their personality structure as the women within the ideal weight range (same mean number of responses). However, the quality of the production between the two groups differed significantly in that the obese women restricted their reactions to a more concrete, a more arid type of association and appeared to function in a more limited sphere (higher "Detail" %).

The obese women have more anxiety than the ideal weights. Because they achieve less mastery over anxiety, they were held in a greater state of tension (higher "Form" %). Although the obese women showed as much emotional rapport with others (no statistical difference between the two groups in the color ratios), they disclosed less pleasure in object relationships (fewer gave "lamp" responses to "phallic" areas, e.g. in cards 6, 9, 1, and 2; also fewer gave "recreation" content). Intolerance of tension and/or acting out did not seem to be a distinguishing factor between the two groups of women (absence of "color" ratio predominance). This suggests that the obese women use a special vehicle for coping with tension.

Repression is deeper in the obese women and hence the unconscious content is less available. There is an increase in pain and decrease in worldly pleasure. There is a higher degree of energy poured into defense and hence more internalization and depressive trends (higher "Form" %). Energies appear to be directed more to the area of internal difficulties, repressed drives and memories (lower number of "Human Movement" responses) than relationship to external objects. The obese women have less flexible ego boundaries. Their reservoir of energy is more tightly bound and not as readily available for constructive use (lower "expanded Form" %). Hence obese women may be expected to react more strongly to traumatic experiences. Their powers of ego integration are less adequate. The obese women are less constructively self-assertive. They resort more to a passive type of mastery and show less operating capacity for achieving active constructive goals. The pattern of the Rorschach variables approximates most nearly the psychic-economic structure of the compulsion neurotic.

An examination of the "trend" results indicated a tendency for the obese women to show less interest in people (less "Human" responses), a tendency for displaying a more blurred sense of reality (more Form-minus responses), and a tendency to present a greater disturbance in body image (more obese did not tend to give "clothing" responses). Of the latter, following Schilder[9], Weiss[11] stated, "It is only when we do not feel at ease in our clothing that we fail to extend to it our bodily ego feeling, instead feel awkward wearing it, as if it were something foreign attached to us."

Summary

Group Rorschach protocols of obese women were contrasted with ideal weight women controls to determine those characteristics which statistically held up through both pilot and experimental groups of 51 and 80 subjects respectively. Those variables which survived the double screening process were then interpretively combined. A personality picture of obese women was constructed that highlighted the characteristics of obese women and which consistently differentiated them from ideal weight women.

Bibliography

1. Abraham, K. *The influence of oral eroticism on character formation.* (*In Selected Papers*) London: Hogarth, 1948.
2. Bruch, H. Psychological aspects of obesity. *Bull. N. Y. Acad. Med.*, 1948, 24, 73-86.
3. Coriat, I. H. Sex and hunger. *Psychoanal. Rev.*, 1921, 8, 375-381.
4. Dublin, L. I., Fisk, E. L., and Kopf, E. W. Physical defects as revealed by periodic health examinations. *Amer. J. Med. Sci.*, 1925, 170, 576-594.
5. Freud, S. *Three contributions to the theory of sex.* Nerv. ment. Dis. Monogr. Ser., New York: Coolidge Foundation, 1910.
6. Kotkov, B. Experiences in group psychotherapy with the obese. *Psychosom. Med.*, in press.
7. Kotkov, B. and Meadow, A. Rorschach criteria for continuing group psychotherapy. *Int. J. Group Psychother.*, in press.
8. Ryder, C. F. and Getting, V. A.: Preliminary report on the health protective clinic. *New Eng. J. Med.*, 1950, 243, 277-280.
9. Schilder, P. *The image and appearance of the human body.* London: Paul, Trench, Trubner, 1935.
10. Sevringhaus, E. L. *Endocrine Therapy in general practice.* Chicago: Year Book Publishers, 1948.
11. Weiss, E. *Principles of psychodynamics.* New York: Grune & Stratton, 1950.
12. Wulff, M. Ueber einem interessanten oralen Symptomenkomplex und seine Beziehung zur Sucht. *Int. Z. Psychoanal.*, 1932, 18, 78-86.

THE DETECTION OF PARANOID TRENDS BY THE EIGHT CARD REDRAWING TEST (8CRT)

LEOPOLD CALIGOR, PH.D.

New York University — Bellevue Medical Center

Introduction

Machover [3], on the basis of clinical experience asserts that paranoid pathology can be detected in figure drawings by (1) "paranoid signs", and by (2) direct inspection. Fisher and Fisher [2], however, tested both of these assumptions, using the drawings of paranoid schizophrenics, and found that "The majority of these drawings did not fall under the category of paranoid in terms of the criteria implicit in each method. Furthermore, it was found that the two separate methods of evaluating paranoid trends did not correlate." The present lack of quantification has been a major obstacle in the further development of the drawing technique as a diagnostic instrument.

The Test

The Eight Card Redrawing Test (8CRT) is believed to open an avenue of approach to quantification—quantification based upon an increase of interrelated content material rather than upon an evaluation of detailed "signs".

The 8CRT consists of eight 8½ x 11 inch sheets of onionskin stapled in a vertical pad-like fashion to a heavy white cardboard base of the same size. A thin sheet of white cardboard of somewhat larger dimensions is inserted between the bottom onionskin and the cardboard. Seven of the eight onionskins are rolled over the top and beneath the cardboard base, with the eighth onionskin resting flat on the white cardboard insert. The instructions for drawing # 1 are: "Make a picture of a whole person." The bottom onionskin is then brought forward from beneath the cardboard base so that it rests flat on the drawing just completed. This presents the subject with a clear tracing image. The instructions for drawing # 2 are: "And now, you may add or subtract anything you want to this picture of a person, or you may change it, or you may leave it alone. Only, make a picture of a whole person." The same procedure and instructions are followed with drawings # 3 through 8. Each time the white cardboard insert is placed beneath the last production made. This prevents the subject from seeing any but the last figure produced, thus making any changes in the latest drawing a development of the immediately preceding one. The average time to administer the 8CRT to adults is about 35 minutes.

The construction of the 8CRT permits observation of the dynamics of change along a continuum. This continuum is believed to be on an increasingly unconscious nature. Graphic aspects which the subject at first feels it is mandatory to include often tend to gradually become excluded, whereas other graphic aspects tend to make a later appearance and to become increasingly reinforced. The graphic changes in the 8CRT frequently provide objectively observable data which lend themselves to content analysis and contribute to insight into the subject's personality dynamics at increasingly unconscious levels.[1]

The goals of the present study are: (1) to compare the ability of the 8CRT and one drawing to detect paranoid pathology, and (2) to determine whether the 8CRT can detect paranoid pathology at increasingly unconscious levels.

Methods

Twenty male subjects, all diagnosed paranoid-schizophrenic on the basis of case history, psychiatric diagnosis and psychological testing, were selected by an

1. A previous study [1] has indicated the increasing ability of later drawings to tap the individual's unconscious conception of his own masculinity-femininity identification.

alphabetic random sampling technique at Brooklyn State Hospital and given the 8CRT. Their drawings #1 and #8 were then removed from the series and each mounted individually.

Three raters who had worked with figure drawings for at least three years, including hospital experience, were shown the forty drawings. Spontaneous notations of the patients were covered up. The raters were instructed: "Here are some figure drawings made by adult males. Your task will be to score them for the presence or absence of strong paranoid trends." Scores were + and −. Each rater scored independently of the other. Where disagreements existed between raters I and II, the evaluation of rater III was used to determine a mutual score.

The twenty 8CRT records were then reassembled to include drawings #1 and #8. The same three raters were now carefully informed of the construction of the 8CRT and instructed: "Your task is to score these series of drawings for the presence or absence of strong paranoid trends. You may use the criteria of content and change in sequence of drawings, breaks in form level, appearance of bizarre specifications, "signs", or any other clinical material made available by the series." The scores were again + and −. As before, where scoring difference existed between raters I and II, the evaluation of rater III was used to determine a mutual score.

Results obtained were used to determine: (1) the efficacy of the 8CRT as compared to that of one drawing in the detection of paranoid pathology by direct clinical inspection, and (2) whether drawings #8 expose paranoid pathology at a more unconscious level than do drawings #1.

Results

Results will be presented in two categories: (1) coarctated schizophrenics (showing primitive, regressed, inflexible drawing characteristics), numbering 9 subjects; (2) over-ideational schizophrenics (showing non-regressed, overactive phantasy, overflexible characteristics), numbering 11 subjects.

Interscorer Agreement. The agreement between raters I and II for the presence or absence of strong paranoid trends for the total sampling (N = 20) was 70% for drawings #1, 80% for drawings #8, and 90% for the 8CRT protocols. The agreement between raters I and II for the coarctated paranoid schizophrenic sampling (N = 9) was 66.7% for drawings #1, 77.8% for drawings #8, and 77.8% for the 8CRT. The agreement between raters I and II for the over-ideational paranoid schizophrenic sampling (N = 11) was 72.7% for drawings #1, 81.8% for drawings #8, and 100% for the 8CRT.

The consistently higher interscorer agreements for the 8CRT than for one drawing suggests the greater stability of the 8CRT as a clinical tool. This is to be expected, since the construction of the 8CRT exposes more graphic data than does a single drawing. This tends to place diagnosis on a scientifically firmer footing and minimizes the subjective "hunch" factor frequently observed in the evaluation of one drawing.

The consistently higher interscorer agreement for drawings #8 than for #1 is believed due to the last drawing of the 8CRT graphically exhibiting more clearly what may be present to a lesser extent on the first one. This is a result of the increasingly unconscious nature of the 8CRT which tends to minimize conscious evasion and to reinforce graphic aspects of which the subject is not aware. The increase in interscorer agreement from drawings #1 to #8 to the 8CRT suggests that although later drawings expose strong paranoid trends to a greater extent than do the earlier ones, the 8CRT reveals even more data than do drawings #8. The difference in increase is believed due primarily to the exposure of inter-related graphic content by the 8CRT which cannot be detected on one drawing.

The consistently higher interscorer agreement for the over-ideational as compared with the coarctated group was expected. Their records tended to expose content, breaks in form level, bizarre specifications, etc., on the later drawings, whereas the coarctated group tended to rigidly reproduce the same figure.

The Detection of Strong Paranoid Trends. The final allocation of strong paranoid trends for drawings #1, #8 and the 8CRT are presented for the coarctated (Table 1a) and overideational paranoid schizophrenic groups (Table 1b) separately.

TABLE 1. FINAL ALLOCATION OF STRONG PARANOID TRENDS ON DRAWINGS NO. 1, NO. 8 AND THE 8CRT FOR A PARANOID SCHIZOPHRENIC SAMPLING

	Ia - Coarctated N = 9				Ib - Ideational N = 11		
Subject	Drawing #1	Drawing #8	8CRT	Subject	Drawing #1	Drawing #8	8CRT
1	—	—	—	10	—	+	+
2	+	+	+	11	—	+	+
3	—	+	+	12	—	—	+
4	—	—	+	13	—	+	+
5	+	+	+	14	—	+	+
6	—	+	+	15	—	+	+
7	+	+	+	16	+	+	+
8	+	+	+	17	—	—	+
9	—	—	—	18	—	+	+
				19	—	—	—
				20	—	+	+
+%	44.4	66.7	77.8	+%	9.1	72.7	90.9

Of the total sampling (N = 20), 25% were scored plus on drawings #1 for the presence of strong paranoid trends, 70% on drawings #8, and 85% on the 8CRT. Of the coarctated paranoid schizophrenic group (N = 9), 44.4% were scored plus on drawings #1, 66.7% on drawings #8, and 77.8% on the 8CRT (Table 1a). Of the over-ideational paranoid schizophrenic group (N = 11), one subject, 9.1%, was scored plus on drawing #1; 8 subjects, 72.7%, on drawing #8; and 10 subjects, 90.9%, on the 8CRT (Table 1b). The only subject scored minus on the 8CRT was scored so by both raters.

For the total sampling, the 8CRT was 3.4 times as good as one drawing in its ability to detect strong paranoid trends. This was due to the 8CRT having frequently exposed interrelated content, overideation, severe breaks in form level, bizarre specifications, paranoid "signs" such as overemphasis on eyes, etc., not observable on drawing #1.

In the coarctated group, 4.9 times as many, were correctly diagnosed for strong paranoid trends on drawings #1 because of their regressed and constricted graphic characteristics. However, the over-ideational protocols for the entire 8CRT permitted the detection of strong paranoid trends 1.2 times as frequently as did the coarctated records for reasons stated above.

The detection of strong paranoid trends on drawings #1 of the over-ideational paranoid schizophrenics was extremely limited, only one subject of 11 having been scored plus. This was due to their successful graphic evasion, maintenance of good form level and the lack of bizarre specifications on one drawing. However, the frequent reversal of these factors and the exposure of inter-related graphic data on the 8CRT made for greater acuity in the detection of strong paranoid trends. This explains why only one subject in his group (9.1%) was scored plus on drawing #1, whereas 10 subjects (90.9%) were scored plus on the entire 8CRT.

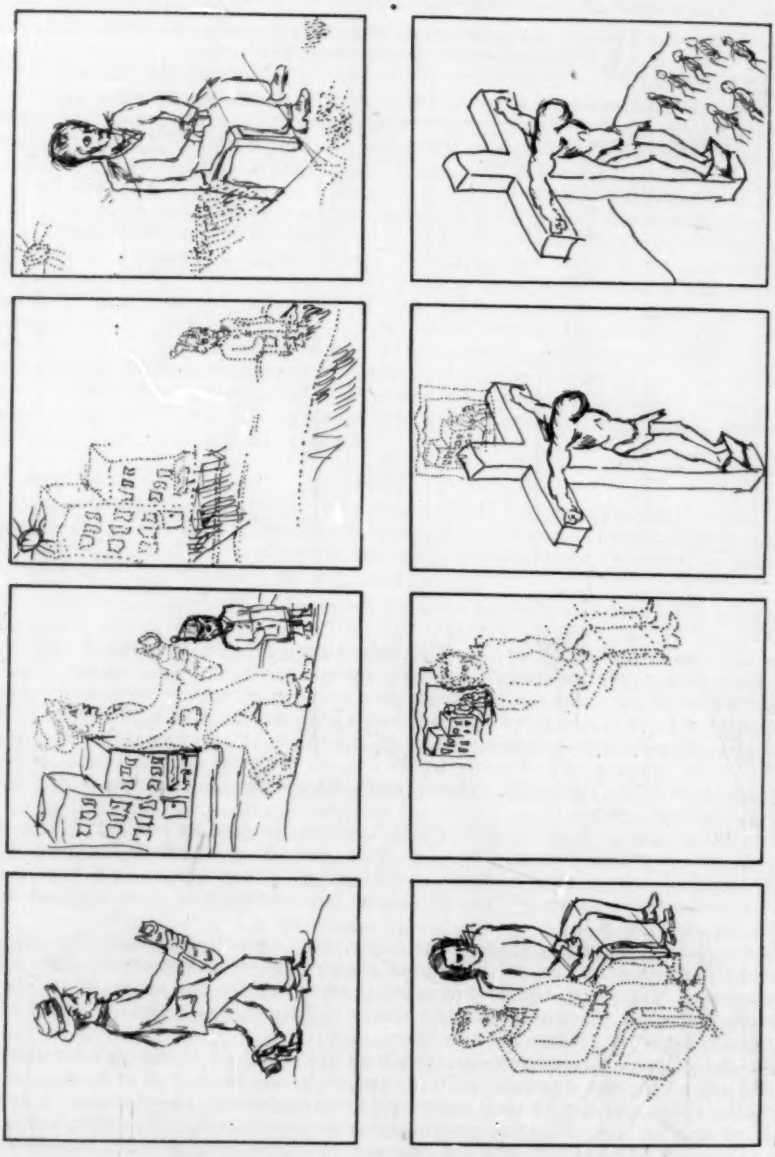

FIG. 1- ILLUSTRATIVE CASE SHOWING SUCCESSIVE DRAWINGS MADE BY A PATIENT ON THE EIGHT CARD REDRAWING TEST (8CRT).

The Detection of Strong Paranoid Trends at Increasingly Unconscious Levels. The increase in the detection of strong paranoid trends of drawings #8 over #1 was 2.8 for the total, 1.5 for the coarctated and 8.0 for the ideational sampling. The consistent increase indicates that later drawings detect strong paranoid trends to a greater extent than do the earlier ones. Due to the 8CRT rationale of the increasingly unconscious nature of later drawings, these results were anticipated; later drawings display graphically more of the true personality dynamics and less of the conscious evasion frequently found on drawings #1. This was obviously more so for the over-ideational than for the coarctated sampling for reasons previously stated.

The detection of strong paranoid trends was consistently greater on the 8CRT than on drawings #8. The increase of the 8CRT over drawings #8 was 15% for the total, 11.1% for the coarctated and 18.2% for the over-ideational sampling. This is due to the exposure of interrelated content which no one drawing could detect and the tapping of increasingly unconscious levels.

Discussion

It must be emphasized that any statistical approach to the 8CRT does the technique injustice as a diagnostic instrument since its major strength is the graphic portrayal of personality dynamics. For example, the over-ideational 8CRT protocols frequently showed evidence of homosexual panic, confusion in masculinity-femininity identification, confusion of sexual and religious symbols, etc. The presence of hostility and paranoid dynamics usually increased on the later drawings. In the coarctated records, once a detail such as an eye was stressed, the emphasis increased until it frequently reached bizarre proportions. The spontaneous verbalizations of the raters indicated that where drawing #1 clearly revealed paranoid symptomatology, the 8CRT was still deemed of value for its portrayal of the underlying psychodynamics.

Summary

The purpose of this paper is to illustrate the eight card method Redrawing Test. It is hoped that the method of the 8CRT will contribute toward the fuller utilization of figure drawings as a diagnostic instrument. Because of the structure of the 8CRT and its completely non-verbal nature which lowers the possibility of evasion, personality dynamics are frequently exposed at increasingly unconscious levels—personality dynamics which are graphically presented by the subject rather than subjectively divined by the clinician. This has been illustrated by a series of 8CRT in the appendix.

Bibliography

1. Caligor, L. The Determination of the Individual's Unconscious Conception of His Own Masculinity-Femininity Identification. *J. Proj. Techs.*, 1951, 15, 494-509.
2. Fisher, S. and Fisher, R. Test of Certain Assumptions Regarding Figure Drawing Analysis. *J. abn. soc. Psychol.*, 1950, 45, 727-732.
3. Machover, K. *Personality projection in the drawing of the human figure.* Springfield, Ill.: Charles C. Thomas, 1949.

THE EFFECTS OF TWO SIMULTANEOUS COGNITIVE AND AFFECTIVE STIMULI

EMANUEL STARER, PH.D.

V. A. Hospital
Coatesville, Pennsylvania

Introduction

This investigation is an exploratory attempt to study the effects of mild affective stimulation on psychomotor performance. Although the approach of employing simultaneous stimuli is not new and had been used in early conditioning experiments, the setting in which the present technique is employed is apparently an innovation. This particular research had been suggested by Malmo, Shagass, Belanger, and Smith whose results led them "to question the neglect of motor aspects of behavior in most studies dealing with personality, and, especially, with its pathological manifestations in psychiatric disorders."[4] Using a control group of 21 normals and 76 psychiatric patients, and a technique of employing a standardized series of painful stimulations (non-specific) they arrived at the conclusion that under this form of non-specific stress, psychoneurotic patients were characterized particularly by disturbances in motor function, this being evidenced by impaired motor control and excessive muscular tension. Their results, in general, supported the hypothesis that psychoneurotic disorders involve defective motor regulation manifested by abnormally increased motor disturbance under any stressful condition.

Method and Procedure

The control group consisted of 15 female affiliate student nurses, age ranged from 19-24, with the mean age of 20, and functioning on at least an average intelligence level. No psychiatric deviations were evident clinically. No difficulty in hearing was present.

The experimental group consisted of 15 female psychiatric patients, all diagnosed as psychotic. The age range consisted of 23-39 years, with the mean of 30.7 years. Three were diagnosed as manic depressive; seven as schizophrenics, paranoid type; three as hebephrenic type; one as catatonic type, and one as mixed schizophrenic type. Two of the patients had improved to the point of being considered ready to leave the hospital, and three others displayed considerable improvement in their behavior. At the time of testing, all were capable of functioning on at least an average intelligence level, as indicated by their performance on the Wechsler-Bellevue Intelligence Scale. No difficulty in hearing was present in any of the patients.

The subjects in both the experimental and control groups were given the Bender Gestalt designs to copy on 8½ in. by 11 in. white paper. They were all provided with two well-sharpened pencils with erasers. The standard instructions were: "Please copy the following designs as well as you can on the paper." If the subjects wanted more paper, they were given it quickly. The time needed for each design was noted.

After the completion of the copying of the Bender Gestalt designs, the subjects were given a five minute rest. Then a recording which had been made on a standard dictaphone machine was played to each of the subjects. It contained the following instructions:

"Please listen to the following instructions. You are going to be given a series of designs to copy as well as you can on the paper. While you are copying them, you will hear me speaking to you. Continue copying the designs no matter what you hear. Do not stop copying the designs at any time. Are you ready? (Pause of 10 seconds). All right, now begin copying the design. (Pause five seconds).

"Now you can hear my voice. Remember to continue copying no matter what you hear. I am going to ask you a series of questions. Do not try to answer them at any time. Just continue copying the designs. (Pause 10 seconds).

1. Do you feel sick at the present time?
2. When you were a child, were you ever happy?
3. Did you sometimes dislike what your mother did?
4. Were your father and mother happy together?
5. Did you ever feel that your father was friendly to you and liked you?
6. Did you feel close to your mother?
7. Did you ever feel jealous of your brothers or sisters or your parents?
8. Did you make a success in your job?
9. Did you ever feel like you wanted to kill yourself?
10. Do you ever get very angry and annoyed with your boy friend or husband?
11. Do you enjoy intercourse or sex activities?
12. Are you happy with your boy friend or your husband?
13. Do you want to get well?
14. Do you feel you can get better?

These questions were spaced 15 seconds apart. If the subject did not complete the Bender Gestalt designs by the time the last question of the series was heard, the questions were again presented from the beginning. During the testing, the subjects were carefully observed at all times. The number of the question which coincided with the different Bender Gestalt designs was noted and recorded.

After the subjects had completed the copying of the designs, accompanied by the added stimulation of affective questions, they were told to write down as many of the questions which they heard that they could remember.

The Bender Gestalt designs were scored for deviations according to the Pascal-Suttell[5] method. In addition, Design A was also scored in a manner similar to that prescribed for the remaining eight designs. The investigator found that this design could not be omitted from the scoring in view of its importance for testing of possible distortions or rotations.

Results

A comparison was made of the results of the experimental and control groups for the following categories: differences in the Bender Gestalt designs before and after the playing of the recording containing affective stimulation; differences in time reaction before and after affective stimulation; the number of questions recalled in each group, and the type of behavior during the testing.

Table 1. Changes in Bender Gestalt Performance as Measured by the Pascal-Suttell Method after Affective Stimulation for Experimental and Control Groups

	Experimental Group N = 15	Control Group N = 15
Improvement	0	5
Same	5	8
Worse	10	2

The results in evaluating the deviations on the Bender Gestalt figures both before and after affective stimulation are shown in table 1. The Chi Square Method shows that the differences between the groups are significant at the one per cent level. It is interesting to note that those cases in the experimental group which showed no change also were considered to be clinically improved, and two of the

patients are considered sufficiently recovered to leave the hospital. The psychomotor disorganization in the experimental group was at times very severe. Rotations, distortions, perseveration, inappropriate planning, and refusing to complete the task were present. It was possible for the investigator to note which of the affective questions was helping to produce this excessive disorganization and to better understand the sensitive areas.

With respect to differences in time reaction in the copying of the Bender Gestalt designs before and after added affective stimulation, it was found that the control group had a record of 11 cases or 73.3 per cent where the time taken on the second performance was either the same or less than the time taken to complete the copying on the first performance. This was markedly different from the result in the experimental group where only four cases or 26.7 per cent were able to complete the task in the same or less time than the initial performance. This again points to the presence of greater blocking in the experimental group under the pressure of affective stress.

TABLE 2. RECALL OF AFFECTIVE STRESS BY EXPERIMENTAL AND CONTROL GROUPS

Groups	Number of Questions Recalled	Mean
Experimental	0 – 10	4.2
Control	4 – 9	6

Differences in the recall of the questions asked during the added affective stimulation performance were noted. The control group had a range of four to nine questions recalled accurately with a mean of six, as compared to the experimental group having a range of zero to ten questions recalled with a mean of 4.2. No distortions were noted in the control group. In the experimental group one case showed extreme distortions in recall and two other cases reported that they couldn't remember any of the questions asked. It was of importance to note that in one of the cases in the experimental group where the psychomotor disorganization was minimal, disorganization was apparent in the recall process. This patient reported that she could not remember any of the questions at all.

With respect to the behavior during the experiment, no unusual or bizarre was noted in the control group. In the experimental group six of the cases showed marked inappropriate behavior. Examples of this included: stopping in the middle of a task coincidental with the presented affective stimuli and displaying extreme inattention, responding to the dictaphone as if it were a real person, inappropriate grinning and smiling when affective stimulation was presented, disturbed looks accompanied by screwing up of the forehead or excessive frowns, and responding to the questions overtly which was contrary to the directions. One patient who could no longer restrain herself said in a very loud voice during the affective stimulation, "Yes, I'm sick, I'm sick." This behavior was not apparent during the initial testing but became easily discernible in six of the cases of the experimental group after the affective stimulation was applied. The results are consistent with those obtained in the Malmo, Shagass, and Belanger Smith[4] study. They tend to suggest that further studies along similar lines are needed.

Although the technique used in the present experiment is subject to possible revision, the basic approach of employing simultaneous impingement of stimuli in a clinical setting is worthy of careful consideration, and merits more extensive research and the author would appreciate the forwarding of any additional data or results of investigations which are the same or similar in design.

SUMMARY

Two groups, a control and experimental, copied the Bender Gestalt designs once and then a second time under the additional stimulation of affective stimuli in the form of personal questions. The control group consisted of 15 affiliate nurse students with a range of 19-24 years and a mean of 20 years. The experimental group consisted of 15 female psychiatric patients, all of whom had been diagnosed as psychotic, with an age range of 23-39 years and a mean of 30.7 years. Differences in the performance between the two groups included psychomotor disorganization, delayed reaction time, some difficulty in recall of the affective stimuli, and behavioral disorganization in the psychiatric group when compared with the control group.

The suggested technique of impingement of simultaneous stimuli, one of which has affective meaning, is worthy of further investigation, particularly since in the clinical setting it may give a clue to underlying dynamics, and serve to indicate how the individual actually reacts to a relatively mild stressful situation.

REFERENCES

1. ERICKSEN, C. W. Perceptual defense as a function of unacceptable needs. *J. abn. soc. Psychol.*, 1951, 46, 557-564.
2. LURIA, A. R. *The nature of human conflicts.* New York: Liverright, 1932.
3. MALMO, R. B. and SHAGASS, C. Physiologic studies of reaction to stress in anxiety and early schizophrenia. *Psychosomat. Med.*, 1949, 11, 9-24.
4. MALMO, R. B., SHAGASS, C., BELANGER, D. J., and SMITH, A. A. Motor control in psychiatric patients under experimental stress. *J. abn. soc. Psychol.*, 1951, 46, 539-547.
5. PASCAL, G. R. and SUTTELL, B. J. *The Bender Gestalt Test.* New York: Grune and Stratton, 1951.

THE TIME-LIMIT VERSION OF THE WORD-ASSOCIATION TEST*

J. P. S. ROBERTSON

Netherne Hospital, Coulsdon, England

INTRODUCTION

A group form of the word-association test had wide use during the Second World War in the selection procedures of the British Army, as described by Harris [2], and Vernon and Parry [6]. More recently the same form has been employed in Britain in the selection procedures of other institutions and has been introduced into some mental hospitals as a diagnostic method in abnormal psychology, as described by Robertson [3]. In this form of the test *S* writes his own responses. He is given a time-limit in which to do so but his *RT* to the stimulus-word is ignored. He is expected to respond not by one word only but by a phrase, a sentence or a series of words. He is also expected to write as much as he can within the time-limit.

This form of the test has come to be known as the *group* word-association test in contrast to the traditional *individual* word-association test from which it was derived. Its essential difference from preceding forms, however, lies not in its being a group test, since group forms of the word-association test have been employed before, but in fact that *S* is given a time-limit and expected to reply by a phrase or sentence, not by a single word. For this reason it can more distinctively be termed the *time-limit*, as opposed to the traditional single-word, version of the word-association test.

COMPARISON OF SINGLE-WORD AND TIME-LIMIT VERSIONS

The circumstance that *S* writes as much as he can within the time-limit transforms the nature of the test. The collection of written responses gives much more

*The author has incorporated some material included in a doctoral thesis accepted by the University of London in 1950.

abundant evidence than do single-word responses about the personality of *S*, based on his chosen modes of formulation and on the attitudes that are elicited from him.

This may be exemplified by quoting blocks of single-word and time-limit responses to the same stimulus-words given by the same psychoneurotic S. The time-limit responses were obtained five weeks after the single-word ones.

Stimulus-word	*One-word response*	*RT*	*Time-Limit response*
TRUNK	Travel	3″	The trunk murders before the war.
DRINK	Beer	3″	I don't like beer.
PARTY	People	4″	I'm shy at parties.
BOOK	Shelf	1″	I read a lot.
LAMP	Light	1″	Streets at night.
RUG	Fire	1″	The rugs at home.
CHAIR	Sedan	1″	A comfortable easy chair.
BOY-FRIEND	Girl-friend	9″	It seems an American word.
DARK	Horse	3″	I used to be afraid of the dark.
DEPRESSED	Feeling	2″	I am often depressed.
SPRING	Flowers	1″	Spring is the best time of the year.
BOWL	Roses	1″	Old men play bowls. My father.
SUIDICE	Death	2″	I would never commit suicide.
MOUNTAIN	Valley	1″	I am interested in geography.
HOUSE	Home	1″	We have got a nice house.

It will be seen that the time-limit responses present a clearer and fuller picture of *S's* attitudes, outlook and modes of reaction. They are more specific, more unambiguous, less superficial and give more scope for the display of individual differences in attitude. Their only disadvantage is in not providing the measure of disturbance given by *RT* in the single-word responses. The difference between the responses obtained by the two methods is not always so marked as in these examples but the trend is unmistakable.

Administration of Time-Limit Version

In order to obtain good results with the time-limit responses careful test instructions are necessary, emphasizing that there are no correct answers and that S can express himself in any way he likes, except for the following.

1. Something must be written for every stimulus-word, leaving no blanks.
2. As much as possible must be written within the time-limit and one-word responses must be avoided.
3. It is not a vocabulary test, so that *S* is not being asked to write down what the words mean.

It is helpful at the beginning to quote examples of possible responses, e. g. "I might give the word DOG. You might write down *My spaniel at home* or *I love them* or *I prefer cats* or *Fox-terriers. Bulldogs. Retrievers. Collies,* or *A dog running after a rabbit* or dozens of other things."

The optimal time-limit is probably 20 seconds but with aged or seriously disordered patients a minute may be necessary and with young, alert *S's* 15 seconds are ample. It is essential to maintain the same time-limit throughout. A visual presentation of the stimulus-words is preferable to an auditory. The test is most successful when administered to a group but can also be applied individually. It is very undesirable in the latter case for the tester to write the responses.

Modes of Individual Difference In Performance

The chief lines along which individual differences appear in the time-limit responses are the following.

1. *The general verbal formulation of responses.* Phrases, sentences, word chains, completions, etc.
2. *The particular verbal formulation of responses.* Queries, exclamations, commands, etc., or the use of negatives, abstract nouns, passive verbs, words such as "must", "ought", "wish", "try", "I", "you", "we", etc.
3. *S's general way of approaching the task set by the test.* Factually, introspectively, valuatively, verbally or irrelevantly. Examples of these are:

(*factual*)	CHILDREN	They spend a lot of time in school.
(*introspective*)	CHILDREN	A feeling that I want to protect them.
(*valuative*)	FATHER	A real good sort. One of the best.
(*verbal*)	FATHER	Eternal Father strong to save.
(*irrelevant*)	FAILURE	Bones in my arm.

4. *The attitudes and opinions elicited by particular stimulus-words.* Elicited attitudes as opposed to formal characteristics depend largely on the word-list used. With mental-hospital patients it has been found useful to employ words covering the following fields: family relationships; social relationships; emotional states; sex and marriage; body parts; disease; death, fear of death and suicide; and miscellaneous trends, such as suspicion, anxiety, aggressiveness, low self-valuation, fussiness, etc.
5. *Special contents and references not elicited by particular stimulus-words.* Allusive responses, quotations, current-event responses, personal-name responses, place-name responses, pre-occupations, etc.

Many individual differences along these five lines appear to be correlated with differences in personality and psychopathology. It is hoped at a future date to publish evidence bearing on this from hundreds of cases tested at Netherne Hospital.

The time-limit test is particularly successful in evoking the deviant modes of verbal expression and thinking characteristic of schizophrenic reactivity. Body words, e.g. BLOOD, STOMACH, etc., are useful in the body-schema aspect of this problem. The following are examples of responses showing the thinking disturbances commonly described as "schizophrenic".

ENVY	Shame with a good heart to it.
ENVY	She is the next sister to jealousy.
BODY	Body odor double fear.
BODY	Skeleton holding a machine.
BLOOD	Red and white corpuscles flow along the mind.
BROTHER	Perhaps my brother is butter, better than margarine.
STOMACH	The pro's and con's of peas and cream.
SUCK	To take in by convulsions and electric fits.
HUSBAND	I believe this is always a male.
ANGER	My brain bursts out of my skull.
MOUTH	The groceries' enemy.

Comparison With Sentence-Completion Tests

The time-limit word-association test has obviously many features in common with sentence-completion tests as described by Shor [5], Rotter and Willerman [4] and Bell [1]. As contrasted with these, it is more effective in disclosing the basic modes of approach which are unwittingly expressed in verbal formulations and in bringing to light thinking disturbances. It is also more speedily administered to disordered persons. It is less successful in pinning the patient down to the expression of an attitude on a particular topic. The two tests, however, conveniently supplement one another when combined in a battery, the time-limit association test being given a group administration and the sentence completion test a more leisured individual administration subsequently.

SUMMARY

A brief account is given of the history, nature and administration of the time-limit version of the word-association test and of the lines along which inferences can be made from it as to personality and psychopathology.

REFERENCES

1. BELL, J. E. *Projective techniques.* New York: Longmans Green, 1948, 45-53.
2. HARRIS, H. *The group approach to leadership-testing.* London: Routledge and Kegan-Paul, 1949, 67-68.
3. ROBERTSON, J. P. S. The group word-association test: its diagnostic capacity in abnormal psychology. Unpublished doctoral thesis, University of London, 1950.
4. ROTTER, J. B. and WILLERMAN, B. The incomplete sentences test. *J. consult. Psychol.*, 1947, *11*, 43-48.
5. SHOR, J. Report on a verbal projective technique. *J. clin. Psychol.*, 1946, *2*, 279-282.
6. VERNON, P. E. and PARRY, J. B. *Personnel selection in the British forces.* London: University of London Press, 1949, 57, 255.

EFFECT OF TEST-RETEST INTERVAL AND ITEM ARRANGEMENT ON THE SHORTENED FORMS OF THE MMPI

GORDON L. MACDONALD

Toledo State Hospital

This study investigated two aspects of test administration connected with the MMPI. One was the effect of the time interval between test and retest, and the other the effect of item arrangement on the individual form of this inventory. It was hypothesized that either, or both, of these might have been responsible for the low statistical values obtained for validity and reliability in an earlier study [1].

PROCEDURE

Groups. There were four groups used in this study: (1) A mixed population of college students and student nurses consisting of 50 males and 67 females, (2) a sub-group of 24 student nurses from the latter population, (3) a group of 10 student nurses, and (4) the control group of 25 student nurses.

Forms. The two forms of the MMPI were shortened by eliminating all of the unscored items including the sixteen duplicated items on the group form. The second appearances of these items were masked out. The scorable items on the group form stop at number 365 with seven items past that number being scored for the K scale. The items on the individual form were arranged in identical order with those on the group form. They were rearranged in this same order after each scoring.

Test Administration. The control group of 25 student nurses took both forms of the test at one session with the forms being alternated in presentation. The items of the individual form were rearranged in the same order as the group form after being scored. The mixed population of 117 and the sub-group of 24 student nurses had a one week interval between tests with the items on the individual form being randomized for each presentation. The group of 10 student nurses took both forms of the test at one sitting with the individual form items being randomized.

Statistical Analysis. Coefficients of correlation for test-retest were computed as were the critical ratios of the differences between the means of test-retest results for the control group of 25 student nurses. The comparisons of the groups were by the critical ratio technique.

RESULTS

Control Group. The coefficients of correlation and the critical ratios obtained for the test-retest comparison of the control group of 25 student nurses are given in

Table 1. The coefficients for the L, F and Mf scales for both raw scores and T-scores are of such magnitude that predictions of individual's scores could be reliably undertaken. The critical ratios of the differences between the means is statistically insignificant for all scales and both scoring methods.

Control Group Compared With Mixed Population. The critical ratios obtained for both forms and both scoring methods for the comparison of the control group with the mixed population of 117 individuals is shown in Table 2. The C. R. for the Mf raw scores, and the K scale T-scores (individual form) are significant at the 1% level. The Hy scale T-scores (individual form) have a C. R. that is significant at the 5% level. The Mf scale raw scores and the K and Mf scales T-scores (group form) have critical ratios that are significant at the 1% level. The K scale raw scores and the Hy and Sc scales T-scores show C. R.'s. significant at the 5% level. All of the other C. R.'s. for the comparisons of these groups are statistically insignificant.

TABLE 1. COEFFICIENTS OF CORRELATION AND CRITICAL RATIOS OF THE DIFFERENCES BETWEEN THE MEANS FOR TEST-RETEST COMPARISON FOR TWENTY-FIVE STUDENT NURSES.

	r Raw Scores	*r* T-Scores	C.R. Raw Scores	C.R. T-Scores
L	.846	.903	.141	.031
F	.847	.878	1.388	.893
K	.795	.677	.067	.172
Hs	.508	.625	.239	.630
D	.844	.570	.370	.211
Hy	.606	.609	.428	.428
Pd	.503	.625	.678	.032
Mf	.894	.895	.595	.629
Pa	.722	.753	.528	.522
Pt	.819	.644	.694	.036
Sc	.847	.562	.368	.584
Ma	.519	.334	.641	.356

TABLE 2. CRITICAL RATIOS OF THE DIFFERENCES BETWEEN THE MEANS COMPARING THE MIXED POPULATION WITH THE CONTROL GROUP.

	Individual Form		Group Form	
	Raw Scores	T-Scores	Raw Scores	T-Scores
L	1.395	.421	1.250	.324
F	.150	1.276	1.105	.146
K	1.729	4.387	2.901	5.068
Hs	.673	1.369	.012	1.942
D	1.054	.167	1.162	.251
Hy	1.770	2.873	2.388	2.854
Pd	.127	.119	.424	.945
Mf	3.087	.408	3.637	.939
Pa	.127	.032	.891	.869
Pt	.358	1.791	.476	1.420
Sc	1.211	1.679	.828	2.536
Ma	1.899	1.874	2.287	5.214

Comparable Groups of Student Nurses. The critical ratios of the differences between the means for both forms and scoring methods derived from comparing the control group with a group of 24 student nurses is shown in Table 3. The C. R.'s. for all scales of both forms and both scoring methods are statistically insignificant.

Time Between Tests Constant. The C. R.'s. of the differences between the means for both forms and scoring methods for the comparison of the control group with a group of ten student nurses is shown in Table 4. Both forms of the test were administered at one session to all subjects in both groups. The C. R.'s. on all scales for both forms and both scoring methods are statistically insignificant.

TABLE 3. CRITICAL RATIOS OF THE DIFFERENCES BETWEEN THE MEANS COMPARING THE SUB-GROUP OF 24 STUDENT NURSES WITH THE CONTROL GROUP

	Individual Form		Group Form	
Scale	Raw Scores	T-Scores	Raw Scores	T-Scores
L	1.840	.045	.173	1.177
F	1.545	1.770	1.708	1.711
K	1.000	.471	.523	.421
Hs	1.345	.210	1.330	.590
D	1.919	2.125	1.743	1.367
Hy	1.371	1.762	1.934	1.520
Pd	2.214	1.776	.384	.430
Mf	.463	.323	.832	.450
Pa	1.229	1.336	.571	.498
Pt	1.807	1.495	2.373	1.460
Sc	1.520	1.075	1.834	.115
Ma	.853	.671	.145	1.654

TABLE 4. CRITICAL RATIOS OF THE DIFFERENCES BETWEEN THE MEANS COMPARING THE GROUP OF 10 STUDENT NURSES WITH THE CONTROL GROUP.

	Individual Form		Group Form	
Scale	Raw Scores	T-Scores	Raw Scores	T-Scores
L	.661	.018	.677	.329
F	.345	.646	1.785	2.073
K	1.034	.646	1.263	.768
Hs	1.794	1.363	1.490	.130
D	1.171	1.322	2.002	1.942
Hy	.512	.577	.193	.219
Pd	.567	.475	1.053	.476
Mf	.189	.143	.618	.706
Pa	.594	.504	1.786	1.853
Pt	.731	.645	2.227	.948
Sc	.520	.484	1.662	.456
Ma	.598	.598	.648	.240

SUMMARY AND CONCLUSIONS

The control group shows less discrepancies in performance from one form to the other, as indicated by the C. R.'s., than did the mixed population in the earlier study [1]. While the coefficients of correlation have been raised in some cases (L, F, Mf scales) to the level where individual predictions could be attempted more reliably, the general improvement in performance has not affected the other scales to the same extent. The results of this comparison of two forms composed of identical items arranged in identical order and administered at the same session leave much to be desired in a test to be used in appraising personality.

The comparison of the control group with the other groups shows that the improvement in test performance of the control group can be considered as statistically insignificant. Those few statistically significant differences that did occur can probably be attributed to differences in performance between the sexes found in the earlier study [1] of the mixed population.

These studies show that item arrangement on the individual form and the time interval between test and retest did not affect performance significantly. The discrepancies of the earlier study cannot be attributed to these factors.

BIBLIOGRAPHY

1. MACDONALD, GORDON L. A Study of the Shortened Group and Individual Forms of the MMPI. *J. clin. Psychol.*, 1952., 8,

THE CONSISTENCY OF LONGITUDINAL PERFORMANCE ON THE GRAHAM-KENDALL MEMORY-FOR-DESIGNS TEST

RENATE GERBOTH ARMSTRONG

Alton State Hospital, Alton, Illinois[1]

Introduction

A pilot study on the Graham-Kendall Memory-for-Designs test was conducted to compare the difference between the performance of a group of schizophrenic and a group of organic patients. The purpose of the study is to determine whether the performance of schizophrenics varies significantly initially and longitudinally from organics. As far as could be determined the only published research on this test was done by its authors several years ago. The test consisted of 15 simple geometric designs, each of which is exposed to the subject for 5 seconds, after which he is asked to reproduce it from memory. Each design is scored on a four point scale ranging from 0 - 3. A score of 0 is given to a satisfactory reproduction; a score of 1 is given when more than 2 easily identifiable errors have been made but the gestalt is retained; a score of 2 is given when the gestalt has been lost; and a score of 3 is given when the design has been reversed or rotated by 90 or 180 degrees.

Procedure

The subjects consisted of 25 organic and 25 schizophrenic Alton State Hospital patients, both groups comprising a rather heterogeneous sample. Since this project is only a preliminary study conducted in a limited period of time, no further variables were taken into account. Each subject was given the test 3 times in about a 10 day period. The materials were then scored and the statistical results evaluated.

Results and Discussion

The schizophrenic group had a mean score for the 3 performances of 7.33 (S.D. 5.13) and the organic patients had a mean score of 14.87 (S.D. 7.29), i.e. the mean for the organics was more than twice as high as that of the schizophrenics. As to be expected then, a much greater amount of deterioration exists in the organic than in the schizophrenic patients. This difference was significant at the 1% level of confidence. Considerable overlap exists between the lower ranges of the schizophrenic group. Individual diagnosis, therefore, can only be completely acceptable at present in the upper one-third of the total test score range. With larger and more homogeneous sampling, finer distinctions may very well be drawn.

Table 1. Comparison of Performance of 25 Organic and 25 Schizophrenic Patients on Three Retests

	Mean	Sigma	Range
Schizophrenic Group	7.33	5.13	0–23
Test 1	8.52	5.26	0–19
2	7.32	6.43	0–23
3	6.16	5.79	0–21
Organic Group	14.87	7.29	0–31
Test 1	16.00	7.83	0–31
2	14.80	8.05	0–30
3	13.80	9.73	0–30

	t	sig.
Difference between Sch. and Org. Means	4.14	.01
Difference between Sch. and Org. Mean Differences	0.88	.50

[1]Now at East Moline State Hospital, Illinois. The author wishes to express her thanks to Melville Finkelstein, Sylvia Finkelstein, and Ruth Westerhold for their help and suggestions.

The statistical results show that within the organic group, Test I is 89% as variable as Test II, but only 69% as variable as Test III. In the schizophrenic group, the percent of variability is much less: Test I is 70% as variable as Test II and 65% as variable as Test III. The organics showed a steady improvement from the first performance to the last. The schizophrenics, on the other hand, gave a much better initial performance but then failed to improve noticeably upon retesting.

The differences between the schizophrenics' M_1–M_2has a t of 0.71, M_1–M_3 has a t of 1.48, and M_2–M_3 has a t of 0.66; the difference between the organics' M_1–M_2 has a t of 0.52, M_1–M_3 has a t of 0.86, and M_2–M_3 has a t of 0.39. The difference between the three means of each group are not statistically significant. However, the schizophrenic group shows more inter-test variability than the organic as seen by the average difference between the individual performance; that of the schizophrenics is 2.23 (S.D. 1.82), and that of the organics only 1.86 (S.D. 0.97). Considering the great differences between the means of the two groups, one would expect the opposite to be in effect.

The standard deviations of the separate means and the combined means of the two groups suggest that the performance of the schizophrenic is by far more erratic than of the organic group. The standard deviations are relatively close, while the range and means vary widely.

Summary and Conclusion

After testing two groups of twenty-five patients three times with the Graham-Kendall Memory-for-Designs Test it was found that the performance of the organic group is, in general, much poorer than that of the schizophrenic group, indicating a more extensive deteriorative process. The overlap between the higher range of the more deteriorated schizophrenics and the lower range of the less deteriorated organics makes individual diagnosis tenuous. However, the retest may clarify this overlap by the amount of variability revealed.

Although we found consistent trends as far as variability of performance is concerned, we cannot make a conclusive statement as to their cause. For example, the apparent improvement upon retesting the organics and their lesser intertest variability could be attributed to several causes. One could assume that these factors may be due to their less erratic behavior as compared to the schizophrenics, or to their greater motivation to improve their performance, and/or due to learning effect. On the other hand, there is a lack of improvement upon retesting in the schizophrenics and a greater intertest variability. With larger and more homogeneous groups further validations may well be obtained, and the underlying reasons and causes for such performance may then be fully exposed and substantiated. It is then that the Graham-Kendall will assume greater value and reliability as a diagnostic tool.

References

1. Graham, F. K. and Kendall, B. S. Performance of Brain Damaged Cases on a Memory-For-Designs Test. *J. Abnorm. & Soc. Psychol.*, 1946, 41, No. 3.
2. Kendall, B. S., and Graham, F. K. Further Standardization of the Memory-For-Designs Test. *J. Consult. Psychol.*, 1948, 12, 349-354.

SUGGESTIONS FOR EFFICIENT PRESENTATION OF THE WECHSLER-BELLEVUE OBJECT-ASSEMBLY SUB-TEST

WALTER SHANNON AND PHILIP D. ROSSI

St. Lawrence University

As every psychologist knows, all of the Wechsler-Bellevue test materials are subject to ease of presentation with the exception of the object-assembly sub-test. The traditional presentation of the object-assembly mutilations has always been awkward, time-consuming, and tending to disrupt rapport. However, by inserting the individual pieces in a specially constructed formboard as illustrated and described in this article, the psychologist will have a quick and efficient means of presentation. A detailed description of how the formboards may be constructed is offered although a study of the photographs in most instances should suffice to demonstrate how construction can be effected.

Figure 1 illustrates how the pieces fit in the formboard already in their proper places. Figure 2 illustrates how the pieces will appear to the subject once they are removed, and also illustrates the appearance of the formboard with pieces removed. Figure 3 illustrates how the panel is removed to allow speedy and accurate arrayment of the mutilations. Note that in this fashion the subject does not see the mutilations until given the instructions to put them together, and that all of the pieces are presented accurately with one manipulation whereas in the traditional

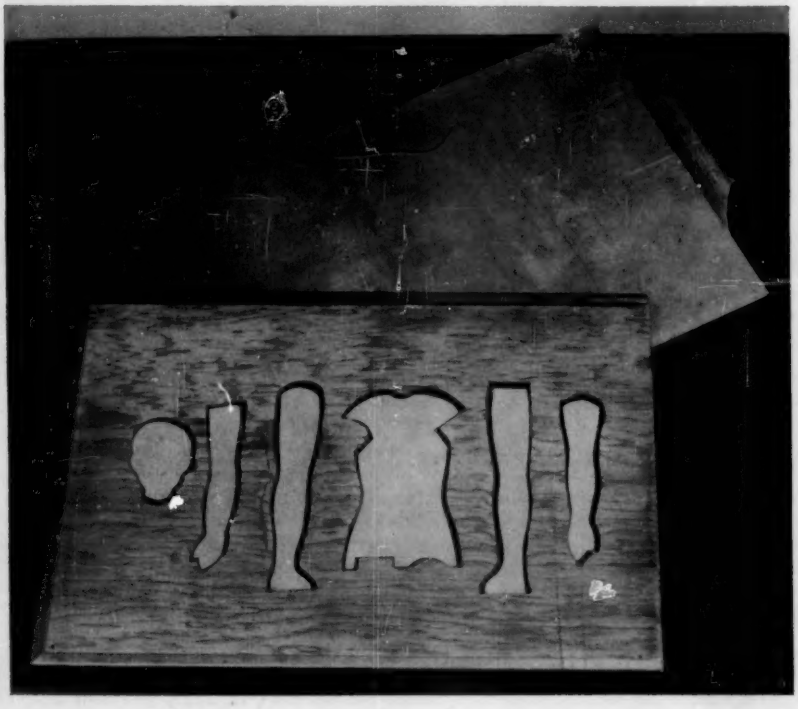

Figure 1

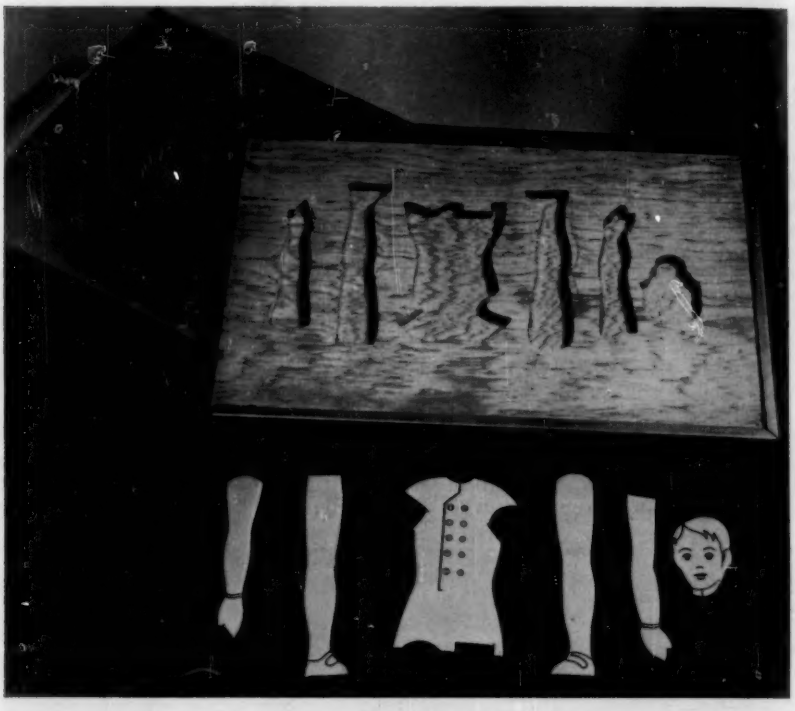

Figure 2

method each piece had to be handled separately. Only photographs of the manikin are offered because the authors felt that illustration of one of the sets would suffice to demonstrate the project.

Construction of the Object-Assembly Formboard

Materials Used. (1) Plywood: Length 14″, Width 9½″, Thickness ½″. (2) Hard Sheet Steel: Thickness 1/32″.

Three Pieces, Hard Sheet Steel, 15½″ x 9 5/16″.
Three Pieces, Sheet Steel Stripping, 37″ x 1½″.

Method of Cutting Figure Outlines in Plywood. Lay out the pieces *face down in reverse order on the plywood* and trace in this fashion. This applies to all three formboards (see Figure 1). This is necessary because the formboard panel which covers the pieces slides from left to right and thereby necessitates a reversal of the position of the board in presentation (see Figures 2 and 3). Cut tracings ¼″ deep and allow a tolerance of at least 1/8″. Sand down edges and rough spots.

Method of Cutting Hard Sheet Steel. In cutting hard sheet steel piece to make the panel, allow 13 13/16″ for the cover and 1 15/16″ for the handle. Cut in two inches on each side at 13 13/16″ length to make the handle (see Figures 1 and 2). The stripping is used on the sides of the plywood to form inserts for the panel sheet.

Figure 3

SUMMARY

In using the formboard presentation of the object-assembly mutilations of the Wechsler-Bellevue Intelligence Scale as described in this article, the authors have found that use of this procedure saves time, eliminates the traditional awkwardness of presentation, insures proper arrangement of the mutilations, and helps to maintain continued rapport in the testing situation.

THE ANCESTRY OF *ID*

PHILIP L. HARRIMAN

Department of Psychology, Bucknell University
Lewisburg, Pennsylvania

Joan Riviere's translation into English of Freud's *The Ego and the Id* reflects the wide connotation of the term *id*. Nevertheless, there is a strong likelihood that an error is being perpetuated with regard to the derivation of this concept. Customarily, the word *id* is said to have been derived from the Latin demonstrative pronoun. Writers who have the background of a classical education point out that the word is the nominative case, singular number, neuter gender of the familiar paradigm *is, ea, id*. In the writings of Cicero *id* denotes "this, that, that thing, that which was previously mentioned." By present-day convention in psychoanalytic writings, *id* is construed to mean the unorganized, inchoate, undisciplined impulses which obey nothing but the laws of elemental cyrenaicism or crass hedonism. The neuter implication is often taken to be appropriate for designating the lack of personal responsibility inherent in the *Id*. This aspect of the connotation seems to be a matter of general agreement among writers. If that were the only important signification of the term, *id* might be regarded as having been derived from the Latin demonstrative pronoun. No question would be raised. There would be, indeed, an etymological consistency between *Id* and *Ego*. *Id*, however, connotes a static entity in classical Latin. Perhaps Groddeck, who suggested the term to Freud, meant to imply a dynamic force. It is very likely that Freud readily accepted the term because *id* has a dynamic signification attributable to some other derivation than the Latin demonstrative pronoun, neuter gender, singular number.

As a matter of fact, the term *id* came into general use among students of biology long before Freud commenced his medical studies. Inevitably, he must have heard the term frequently mentioned in university lectures dealing with the Weismannian revision of the "Lamarckian error." During the years 1868-1876, August Weismann published a notable series of papers dealing with the mechanisms of heredity and with other major issues in biological theory. His sight having been impaired, he was forced to give up microscopical work and to devote himself to theoretical formulations. His influence was great throughout the university world, and he was generally regarded as the greatest authority on the subject of biology. Naturally, therefore, Groddeck and Freud heard much about the Weismannian concepts during their student days. It is likely that Freud gave ready acceptance to the use of *id* as an appropriate concept in psychoanalytic theory because he found it to connote a dynamic concept. If this conjecture have any validity, he did not intend to imply that he took the word from the familiar Latin sources. On the contrary, *id* is used by Weismann to denote something like the *gene* in contemporary biological writings.

The student will recall that Weismann postulated the existence of *biophores* as the ultimate supramolecular vital unit. They, in turn, are organized into *determinants*. Both *biophores* and *determinants* are submicroscopic in size. Both of them exert dynamic influences upon the course and the nature of development. The *determinants* are clustered in the visible granules of the chromatin, and these he termed *ids*. The chromosomes he considered to be made up of still larger units termed *idants*. Both *ids* and *idants* are to be observed through the microscope. Hence, they are not hypothetical constructs but observable entities. The *id*, therefore, is the smallest observable determinant of those biological processes subsumed under the terms *growth* and *development*. The effect of *ids*, of course, is dynamic and interactive. *Ids*, in a sense, represent tendencies for development; and out of them arise all the differentiated structures of the mature organism. According to Weismann, the growth process is the very antithesis of a static procedure that would lend itself to symbolization by a demonstrative pronoun in the neuter gender. Quite to the contrary, the Weismannian concept of *ids* may be logically extended into psychological

theory of the dynamic type. The basic point of view in Weismann is logically consistent with the dynamic principle of entelechy in the writings of Aristotle. It is very plausible to believe that Freud was familiar with these concepts about the dynamic nature of hereditary determination of organisms. In fact, it is doubtful whether any European university student of biology could have failed to master Weismann's theories and to acquire his terminology. They constituted a forceful negation to all views previously upheld in medical schools. Particularly were they readily accessible to students in German-speaking countries. Freud could not have remained uninfluenced by this theory of biological determinants in growth and development. He must have read and heard much about *ids* during the early years while he was a medical student and a research investigator in neurology.

The derivation of *id* from Weismannian theory does appear to be in close accord with its connotation in psychoanalytic theory. Were it to be referred to classical Latin, Freud might have justified his choice of the term. As a matter of fact, he tacitly assumed that *id* does have an appropriate dynamic connotation which is useful both in accounting for biological and for psychological development. Of course, he invested the term with much broader significations than it ever had in biological theory. Perhaps his choice of *id* was a logical, meaningful extension of a word he had often encountered in earlier years. Certainly, both in Weismann's writings and in psychoanalytic theory *id* has a dynamic connotation. The Latin demonstrative pronoun, on the other hand, has a rather meaningless, neutral, static value if used to symbolize this Freudian referrent.

NOTE ON A HORNEY HYPOTHESIS

WILLIAM SEEMAN

Section of Neurology and Psychiatry,
Mayo Clinic, Rochester, Minnesota

Introduction

The problem with which this paper is concerned is Horney's hypothesis which states that "the basis of anxiety usually lies . . . in hostile impulses . . . " (2, p. 63) For our purposes this may be stated in correlation language as a hypothesis asserting a significant association between hostility and anxiety. That Horney views her conceptual analysis of anxiety as being different from Freud's analysis is suggested by her statement that ". . . the frequency with which anxiety is generated by sexual impulses is largely dependent on the existing cultural attitude towards sexuality. I do not see that sexuality as such is a specific source of anxiety . . . A third point in which I find myself at variance with Freud is his assumption that anxiety is generated only in childhood . . . and that anxiety occurring later in life is based on reactions which have remained infantile". (2, pp. 76-77)

Although Horney nowhere presents evidence relevant to the validity of this formulation, it is an intriguing and plausible hypothesis. The present note is primarily intended to illustrate one way in which the theoretical hypothesis can be given empirical meaning in terms of a logically entailed deduction. If it were shown that a specific daydream could be established in a statistically reliable fashion as a "hostile" daydream, then, according to Horney's formulation, there should be a significantly covarying anxiety.

Method and Results

Five daydreams known to be widely experienced (1, 3) were selected. These were submitted to six judges (two psychoanalysts, two psychiatrists and two clinical psychologists) who were asked to select the daydream regarded as most hostile. The daydreams are as follows:

Type 1: In this daydream you visualize yourself as performing or having performed some daring or heroic act or feat, such as rescuing a person from danger or certain death, or facing some tremendous danger yourself and overcoming it. As a result of this heroic act you are publicly admired and acclaimed, and your friends heap admiration on you.

Type 2: In this daydream you visualize yourself as being extremely attractive. You are complimented on your appearance and your clothes. At gatherings and social functions you are the center of attraction, receiving attentions from many persons present, especially members of the opposite sex.

Type 3: In this daydream you visualize yourself as being promoted in your job to a position of greater importance as well as greater remuneration. You may visualize yourself as directing activities, giving orders, and so forth. You may see yourself at a banquet in your honor with officials making speeches about your abilities.

Type 4: In this daydream you visualize yourself as greatly wronged or misunderstood by friends or parents or both. You visualize yourself as meeting some great misfortune, perhaps even death. When your friends or parents discover that this great misfortune has occurred, they are sorry, even miserable.

Type 5: In this daydream you visualize yourself talking back to, or winning an argument against a person to whom in real life you would *like* to talk back or with whom you would like to argue, but somehow feel that you never can or dare.

TABLE 1. ANXIETY ASSOCIATED WITH FIVE DAYDREAM TYPES AS REPORTED BY THREE SAMPLES OF COLLEGE STUDENTS

Day-dream	Minnesota 1				Minnesota 2				Ohio State			
	Anxiety		No anxiety		Anxiety		No anxiety		Anxiety		No anxiety	
	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent	Number	Per cent
1	56	32.7	115	67.3	34	35.8	61	64.2	15	21.1	56	78.9
2	38	29.0	93	71.0	18	25.7	52	74.3	21	32.8	43	67.2
3	41	27.3	109	72.7	42	38.2	68	61.8	14	24.6	43	75.4
4	103	65.2	55	34.8	54	69.2	24	30.8	42	62.7	25	37.3
5	95	49.7	96	50.3	39	40.2	58	59.8	28	38.4	45	61.6

Horney's hypothesis lays down the following theoretical requirement: That daydream which is most heavily loaded with hostility (by the criterion of statistically reliable agreement of ratings by trained and competent individuals) should elicit anxiety with a significantly greater frequency than do the other daydreams[1]. The ratings with regard to hostility turned out to be unanimously in agreement, which places the reliability of the criterion well beyond the 0.01 level of confidence. The relevant data reporting anxiety associated with each of the daydream types are presented in table 1. Three samples of college students were asked to rate the degree of anxiety associated with each of the daydream types; no anxiety was indicated by a rating of "0", some degree of anxiety by a rating of "1," and considerable anxiety by a rating of "2." For purposes of economy the "1" and "2" ratings are combined in the table under the columns marked "Anxiety." The differences between anxiety ratings associated with daydream 4, that rated unanimously as "most hostile" by the judges, and the anxiety ratings associated with the other daydream types are all significant beyond the 0.01 level; there are twelve such statistically reliable differences, four for each sample. These are presented in table 2.

TABLE 2. DIFFERENCES IN DEGREE OF ANXIETY ASSOCIATED WITH "HOSTILE" DAYDREAM AND ANXIETY ASSOCIATED WITH OTHER DAYDREAMS IN THREE SAMPLES OF COLLEGE STUDENTS*

Sample	Difference between daydream 4 and daydream			
	1	2	3	5
Minnesota 1	32.5	36.2	37.9	15.5
Minnesota 2	33.4	43.5	31.0	29.0
Ohio State	41.6	29.9	38.1	24.3

*Differences of the percentages given in table 1; all are significant beyond the 0.01 level.

SUMMARY

A Horney hypothesis states that hostility and anxiety will be found to covary significantly. This paper has presented evidence to the effect that anxiety associated with the one of five daydreams rated as "most hostile" (with agreement of six judges beyond the 0.01 level) is reliably greater than chance in three samples of college students. These differences are in the direction theoretically required by the hypothesis and may thus be held to be relevant to its validity.

REFERENCES

1. GREEN, G. H. *Psychoanalysis in the Classroom.* New York: G. P. Putnam's Sons, 1922, 272 pp.
2. HORNEY, KAREN. *The Neurotic Personality of Our Time.* New York: W. W. Norton & Company, Inc., 1937, 299 pp.
3. SHAFFER, L. F. *The Psychology of Adjustment: An Objective Approach to Mental Hygiene.* London: George G. Harrap & Co., Ltd., 1936, 622 pp.

EDITORIAL OPINION

REFLECTIONS ON EDITORIAL POLICIES

With this issue marking the end of eight years of publication of this Journal, it seems desirable to review the general objectives and policies which have controlled our actions. This attempt at self-evaluation is stimulated partially by several evaluative reviews of the Journal's contents from external observers but more particularly by our own editorial awareness of our responsibilities for periodically evaluating and publicizing our policies. Every editor has large responsibilities in his role of steward of a publication enterprise serving the interests of science and public welfare. It is almost impossible to administer publication policies in a completely neutral, unbiassed and objective manner, since with such a large supply of manuscripts offered for publication, an important act of decision must be made in the acceptance or rejection of every single paper. As experiences with publication practices in authoritarian states have demonstrated, it is entirely possible for editors to subtlely control the directions of development of a science simply by their selection practices. This can be accomplished by simply accepting papers which happen to be consistent with editorial biases and rejecting all others. This issue becomes pointed up more sharply when a publication adopts an active editorial program which frankly seeks to influence opinion and secure objectives. Assumption of such a role of leadership places obligations and responsibilities upon the editor and his editorial board to make certain that the directions of these attempts at leadership are healthy, progressive, advanced for their times, and genuinely constructive for the advancement of science and social welfare. Undoubtedly the actual accomplishments of any editor will inevitably fall short of attainment of these laudable ideals because of human frailties and situational handicaps which can never be completely compensated for.

The following evaluations may well be prefaced by a brief statement from your editor concerning his own biases and deliberate intentions concerning editorial policy. We recognize in ourselves a strong and perhaps constitutional inclination to *action*, believing that it is preferable to make mistakes in an active program than to do nothing at all. We also perceive in ourselves a strong belief that human affairs can be influenced (even though to small degrees in some situations) by rational leadership which is capable of validly discerning or predicting future trends 5, 10 or even 20 years ahead of current public opinion. Finally, we feel that an editor has a strong obligation to call shots as he sees them, even at the expense of coming in conflict with different schools of thought or the interests of local pressure groups. In the history of human thinking, no great advance has been accomplished without arousing embittered opposition from those whose ideological position is challenged. An editor must choose between a fleeting popularity achieved by conformance with the views of contemporary "leaders", or with the more remote possibility of making a more lasting contribution achieved in the face of opposition from entrenched interests. Contemporary estimates concerning the ultimate value of such potential contributions are largely worthless. Only the test of time can yield ultimate proof of their validity.

In 1945, at the time of beginning of publication of this Journal, the status of psychological science was judged by many observers (and particularly by physical scientists) to be almost chaotic. Not only was there a babel of conflicting schools clamoring to be heard, but there was also a complete lack of administrative or promotional organization in the field. Suddenly during and after the period of World War II, the infant science of clinical psychology had huge social demands and responsibilities placed upon it by an enlightened public demand for more psychological

services. Was this growing-up period of clinical psychology under conditions of great stress to be accomplished by trial-and-error behavior evolving slowly over a long period of time, or was it possible to telescope the process by the intensive application of insightful learning derived from study of the development of older sciences which had already outlived their growing pains? Was it not possible to save many years of experimentation and countless errors by learning from the experience of others? It was within this framework of events that we reached an individual decision to launch a new independent Journal which would serve as a loyal opposition and "needler" to the necessarily more conservative official organizations and publications.

The launching of the project in 1945 under wartime condition with private financing was accomplished only under great and unforeseen difficulties. The greatest difficulty was in securing sufficient papers to publish the first issue (and even the first three volumes) on schedule. For example, with the first issue and later on several occasions, we had sufficient papers promised or scheduled to get the issue out, only to have insufficient material arrive in publishable condition. In spité of the contention voiced by some groups that insufficient publication outlets exist in clinical psychology, the fact seems to be that ample facilities exist to publish all the good papers (and also a large quota of not so good papers) providing that they are edited so as to avoid padding and undue theorizing. In any case, your editor received during the first years of publication a large number of second-rate papers many of which were tattered and torn dust-stained manuscripts which had withstood the creasings of many previous editorial rejections but which were still being submitted by parents who hoped that sometime somewhere somebody might accept them for publication. Under these conditions (and also because we believed that many things needed to be brought out in print), your editor arranged several symposia and monographs and also included a series of papers on directive psychotherapy and other topics which seemed important to us. Although we have recognized the validity of the advice once given us by the late Willard Valentine not to publish our own material, many of the early issues would never have appeared on schedule except by addition of such material to make up the duly advertised number of pages. Fortunately, this condition has remedied itself in the last three years after a reliable flow of good manuscripts finally established itself. Operating under such conditions, it was inevitable that some sacrificing of editorial standards and principles was necessary simply to accomplish the practical matter of getting the Journal out so that the good papers would be published and our obligations to subscribers and others satisfied. That this was accomplished to a satisfactory degree is partially attested to by steadily increasing influence and subscription lists.

A second and continuing source of difficulty has been the problem of securing editorial board participation in an *active* shaping of policy and editorial contribution. While the editing of a scientific journal should theoretically be a group endeavor, it is in practice very difficult to secure more active participation from board members than occasionally submitting or evaluating manuscripts. The editorial pages of this Journal have always been open to any board member (and also on occasion to unsolicited contributors) but only on a few occasions have any of them felt inspired to express themselves. The board has made one important contribution, however, in evaluating and criticizing all editorial material before it is published, an agreement having been made that nothing will be published except with the unanimous approval as to general acceptability by the entire board. Having adopted an active editorial policy, we would like nothing better than to run an open forum of editorial opinion, but failing that, there is nothing left but to continue the department alone or to abandon it completely. The latter step has been considered on several occasions, but the steady flow of commendatory letters from readers has persuaded us to continue it.

A third problem has involved the evolution of editorial standards both for the acceptance of manuscripts and for their editing according to the style of the Journal.

Selection problems are influenced by our opinion that there is currently only a very limited amount of high-grade research being performed in clinical psychology. Although the calibre and quantity of research is slowly improving, only a small number of really high grade manuscripts are submitted. Although the official Journals have slightly better material to choose from, most editors receive a few outstanding manuscripts in each quarter and must fill out their issues with secondary material of more limited interest and research significance. The best an editor can do therefore is to publish a representative cross-section of what he is currently receiving while at the same time stimulating and soliciting the type of manuscripts which he would really like to receive. We have made special efforts to keep this Journal stimulating and readable not only by manuscript selection policies but also by requiring authors to (a) condense material to a brief statement of method and results, (b) to simplify textual, tabular and graphic presentations of data by avoiding undue duplication, (c) by eliminating over-extensive statistical data due to the fact that we assume that contributors are sophisticated enough to make their calculations correctly, (d) by rigorously deleting imaginative theorizing, and (e) by attempting to be genuinely eclectic and thereby avoiding undue adherence to any school of thought. We have insisted on brief concise research reports because we believe that our readers are interested in a wide coverage of all important trends in the field rather than in long detailed reports from a few authors. In our opinion, there is very little research being done in the field which cannot be fairly encompassed in four or five printed pages.

A statistical analysis of evaluative correspondence received from our readers since 1945 would result in a bimodal distribution of opinions skewed more strongly on the side of those who commend our efforts. As would be expected, our editorial opinions have stimulated the most numerous reactions since many of them have been of admittedly controversial nature. We have a vociferous group of loyal adherents who appear to support our ideas wholeheartedly, and on the other hand we have numerous critics some of whom are so emotionally involved and personally antagonistic as to discontinue their subscriptions and express book-burner sentiments. We regret the latter attitudes since it also reflects on our innocent contributors who have had nothing to do with the controversial editorial opinions. Such a situation could only have been anticipated in relation to an editorial policy which has dealt with as many controversial topics as we have. In partial extenuation, we wish to assure our readers that our motives have been the best even though at times our judgment may have erred. In our own professional development, we have been perennially embarrassed over the falibility of attitudes which we held so true only last year. We are not satisfied with ourself or with the state of perfection of the Journal. But we continue to derive some solace over having been the means of bringing many ideas to the attention of the profession, and we will be satisfied if only fifty-one per cent of them bear some ultimate fruit.

F. C. T.

BOOK REVIEWS

Symposium. *The Biology of Mental Health and Disease.* New York: Paul B. Hoeber, 1952. Pp. 684. $10.00.

These are the proceedings of the 27th Annual Conference of the Milbank Memorial Fund which was held in November 1950 at the New York Academy of Medicine. Thirty-eight principal papers are included with commentaries by 108 contributors representing all of the basic and applied fields in psychological science. All of the papers in this volume are so authoritative and up-to-date that it will serve as an important reference work for advanced students for some years to come.

Ferguson, Leonard W. *Personality Measurement.* New York: McGraw-Hill, 1952. Pp. 457. $6.00.

The author is associated with the Aetna Life Insurance and affiliated Companies. This is an authoritative textbook based solidly on the experimental-statistical approach to personality measurement. Its organization is markedly different from older books in the field in that it places greater emphasis on measures of interests and attitudes, adjustment and performance rather than upon psychometrics.

Arthur, Grace. *The Arthur Adaptation of the Leiter International Performance Scale.* Washington: Psychological Service Center Press, 1952. Pp. 73.

This is a manual outlining the rationale and directions for administering the Arthur adaptation of the Leiter Scale. Detailed scoring instructions with illustrations are given for each age level.

Wolff, Charlotte. *The Hand in Psychological Diagnosis.* New York: Philosophical Library, 1952. Pp. 218. $7.50.

Dr. Wolff is a fellow of the British Psychological Society who has been making comparative morphological and psychological studies of the human hand for more than 15 years. This book presents a large amount of statistical data which she has collected concerning variations in the structure of the hand occurring in various constitutional types, in normal persons, in mental deficiency and in mental illness. Included are measurements and observations derived from general inspection, measurement of the hands in relation to height and weight, endocrine factors, examination of fingers and nails, and inspection of the creases of the palm. Although the fields of palmistry and chirognomy are currently in scientific disrepute, this book represents a serious scientific attempt to collect and correlate some long-known observations concerning the hand in such conditions as mongolism.

Kelly, E. Lowell and Fiske, Donald W. *The Prediction of Performance in Clinical Psychology.* Ann Arbor: University of Michigan Press, 1951. Pp. 310.

This is the report of a five year research program sponsored by the Veterans Administration and directed at the evaluation of techniques for the selection of professional personnel in the field of clinical psychology. The scope of the research was greater than any previously attempted and resulted in some very interesting findings comparing clinical and objective approaches to prediction. Every clinical psychologist should study the methods and conclusions reached in this report.

Jershild, Arthur T. *In Search of Self.* New York: Bureau of Publications, Teachers College, 1952. Pp. 147. $2.75.

This little volume by a professor of education at Teachers College is directed toward the teaching profession in encouraging them to explore the role of the school in promoting self-understanding. It outlines how the teacher may go about elementary counseling activities in studying the child in order to help him understand himself.

SHOSTROM, EVERETT L. and BRAMMER, LAWRENCE M. *The Dynamics of the Counseling Process.* New York: McGraw-Hill, 1952. Pp. 213. $3.50.

Dr. Shostrom is head of the Psychology Department at George Pepperdine College, and Dr. Brammer is Guidance Officer and assistant professor of psychology at Sacramento State College. This book outlines an eclectic approach called *self-adjustive* counseling which seeks to aid the client in becoming more self-directive and self-responsible. To assist the client to alter his self-concepts in more mature healthy directions, the authors propose a combination of nondirective and directive methods. This book is oriented primarily to counseling under educational auspices, and most of the techniques are illustrated with materials from high school or college level counseling. The primary emphasis is on rather superficial level educational, vocational and educational problems with little or no attempt to deal with depth problems.

BYCHOWSKI, GUSTAV and DESPERT, J. LOUISE (Eds). *Specialized Techniques in Psychotherapy.* New York: Basic Books, 1952. Pp. 371. $5.00.

The Psychiatric Forum Group consists of an informal monthly gathering of psychiatrists and workers in related field meeting in New York City to interchange the latest approaches in the field. Seventeen distinguished authors have contributed papers on specialized techniques in psychotherapy, including HOCH and POLATIN on Narcodiagnosis and Narcotherapy, LINDER on Hypnoanalysis as a psychotherapeutic technique, EISENBUD on the Telepathy Hypothesis in Psychotherapy, MAX M. STERN on Free Painting as an Auxiliary Technique, SPOTNITZ on group therapy, MITTELMANN on Simultaneous Treatment of Both Parents and Child, BEATA RANK on the psychoanalysis of young children, DESPERT on child schizophrenia, FROMM-REICHMANN on the psychoanalytic treatment of schizophrenia, BYCHOWSKI on psychotherapy with the crippled and disabled, I. PETER GLAUBER on dynamic therapy for the stutterer, RUTH FOX on psychotherapeutics of alcoholism, DAVID ABRAHAMSEN on the treatment of sex offenders, MELITA SPERLING on treatment in psychosomatic medicine, VICTOR W. EISENSTEIN on the differential psychotherapy of borderline states, BELLAK on the emergency treatment of depressions, and PIOTROWSKI and SCHREIBER on Rorschach changes during analytic therapy.

AUSUBEL, DAVID P. *Ego Development and the Personality Disorders.* New York: Grune and Stratton, 1952, Pp. 564. $10.00.

The author is associated with the Bureau of Research and Service of the College of Education, University of Illinois. The book consists of a series of essays presenting the historical development of ego psychology, its dynamics, crises and failures and failures of ego development, the contribution of parent-child relationships to ego development, the relation of anxiety, the development of conscience, and a developmental approach to psychopathology. The concept of *satellization* is introduced to describe a process of reorganization of ego structure whereby the child identifies with dominant parental figures in order to compromise between the two unfavorable alternatives of maintaining ideas of omnipotence or of grossly devaluating the self to meet the demands of reality. Various consequences of healthy satellization are discussed on theoretical levels, together with examples of failure of satellization.

AMES, LOUISE B.; LEARNED, JANET; METRAUX, RUTH W.; and WALKER, RICHARD N. *Child Rorschach Responses.* New York: Paul B. Hoeber, 1952. Pp. 310. $7.50.

This important book from the Gesell Institute of Child Development presents a comprehensive longitudinal study of Rorschach responses from ages two to ten years based on findings from a group of 650 children. Part one involving the first third of the book presents a review of the literature, outline of research methods and scoring, and discussion of the various determinants. The latter two thirds are given over to detailed presentations of typical patterns and illustrative cases at each age level. This source book is a most important milestone in the objectification of Rorschach responses.

ALEXANDER, FRANZ and ROSS, HELEN, (Eds.). *Dynamic Psychiatry.* Chicago: University of Chicago Press, 1952. Pp. 578. $10.00.

Fifteen leaders, including 13 psychiatrists, 1 clinical psychologist and 1 anthropologist, have contributed papers on various theoretical and clinical aspects of dynamic psychiatry. This book continues the current tradition which identifies "dynamic" thinking in psychiatry with psychoanalysis. Within this framework, this book is an authoritative summary of current opinion among analytically oriented workers.

BYCHOWSKI, GUSTAV. *Psychotherapy of Psychosis.* New York: Grune & Stratton, 1952. Pp. 328. $5.75.

The author is assistant clinical professor of Psychiatry at the New York University College of Medicine. It consists of 34 chapters discussing his experiences in dealing with various types of psychotic reaction. Its most important contribution is the demonstration that many psychoses involve personality reactions which can be dealt with by psychotherapy.

GORLOW, LEON; HOCH, ERASMUS, L.; and TELSCHOW, EARL F. *The Nature of Nondirective Group Therapy.* New York: Bureau of Publications, Teachers College, 1952. Pp. 143. $3.25.

This monograph presents a research evaluation of nondirective group therapy as applied to a group of 17 graduate students at Teachers College who requested help with personal problems. In spite of the smallness of the sample, the results of the project are significant because of the thoroughness of the statistical analysis. Special emphasis was given to the nature of the group process, the analysis of the behavior of the members as therapists, and of the role of the leader.

KRIS, ERNST. *Psychoanalytic Explorations in Art.* New York: International Universities Press, 1952. Pp. 358. $7.50.

Dr. Kriss is associated with the Child Study Center of the Yale University School of Medicine. In this book he presents the results of more than 25 years of study in the field of the psychology of art in the form of a series of stimulating essays. Part one deals with the history of art with special emphasis on the potential contributions of psychoanalysis. Part two deals with the art of the insane with detailed consideration of the cases of a psychotic sculptor of the 18th century and of a schizophrenic artist. Part three presents some observations on the psychology of comic art and caricature, with comments on laughter as an expressive process. Part four discusses problems of literary criticism, and a concluding section of two chapters offers reflections on the psychology of creative processes. There is an appendix with 79 illustrations of artistic productions which are commented upon in the text.

FROSCH, JOHN (Ed.) *The Annual Survey of Psychoanalysis. Vol. I.* New York: International Universities Press, 1952. Pp. 556. $10.00.

This is the first of a proposed series of annual surveys of psychoanalysis contributed by leading authorities in the field who present detailed discussions of the most significant papers and books which appeared during 1950.

WOLFF, WERNER. *The Dream—Mirror of Conscience.* New York: Grune and Stratton, 1952, Pp. 348. $8.50.

Dr. Wolff is professor of psychology at Bard College and a prolific writer in the field of abnormal psychology. In this book he reviews the historical development of dream psychology, summarizes the most significant theoretical viewpoints with special emphasis on the contributions of psychoanalysis, and presents an interesting series of illustrations of typical dream patterns.

VINACKE, W. EDGAR. *The Psychology of Thinking.* New York: McGraw-Hill, 1952. Pp. 392. $5.50.

The author is assistant professor of psychology at the University of Hawaii. The book is an experimentally oriented, systematic study of the psychology of thinking covering the historical and theoretical background of the field, reasoning, imagination and personalized thinking.

Attention Subscribers:

Complete your back files.

We can provide extra copies of almost all back issues except January 1945 and January 1948 at a cost of only $2.00 per copy.

Order missing back numbers now.

$2.00 per copy.

Journal of Clinical Psychology
5 Pearl Street Brandon, Vermont

CLINICAL STAFF

MEDICAL STAFF OF PENNSYLVANIA

Robert Devereux, M.D.
Ruth E. Duffy, M.D.
Herbert H. Herskovitz, M.D.
Joseph J. Peters, M.D.
Calvin F. Settlage, M.D.

PSYCHOLOGICAL STAFF OF PENNSYLVANIA

Edgar A. Doll, Ph.D.
Director of Research
Milton Brutten, Ph.D.
Michael B. Dunn, A.M.
Robert G. Ferguson, A.M.
Edward L. French, Ph.D.
Marguerite B. Horn, A.M.
John R. Kleiser, A.M.
Kathryn Kramer
Mary J. Pawling, A.M.
Jack Shelley, M.Ed.

PROFESSIONAL STAFF, THE DEVEREUX RANCH SCHOOL, CALIFORNIA

Charles M. Campbell, Jr., M.D.
Consulting Pediatrician
Richard H. Lambert, M.D.
Consulting Psychiatrist
Ivan A. McGuire, M.D.
Consulting Psychiatrist
David L. Reeves, M.D.
Consulting Neurologist

Robert L. Brigden, Ph.D.
Director of Ranch School
Thomas W. Jefferson, Ph.D.
Clinical Psychologist

ACHIEVEMENT
in Intellectual Growth through Therapeutic Guidan

For the acutely disturbed child wh unresolved conflicts are expressed sym tomatically as an inhibition of intellect activity, Devereux Schools offer a compl program of academic, vocational, psyc logical, and psychiatric guidance. T coordinated efforts of the entire staff p vide each child with the specific help needs to achieve his maximum pote growth--both emotionally and intell ually.

When you encounter a school-aged ch whose normal intellectual capacity is p chically limited by emotional disturban you are invited to let us evaluate the tential outcome of Devereux' specializ education with therapy. Our experien staff will thoroughly review each case h tory and offer a detailed report.

Please address your inquiries to
JOHN M. BARCLAY, *Registrar*

DEVEREUX FOUNDATION
CALIFORNIA PENNSYLVANIA

Devereux Schools

HELENA T. DEVEREUX, *Director*
J. CLIFFORD SCOTT, M.D., *Executive Director*
SANTA BARBARA, CALIFORNIA • DEVON, PENNSYLVA